VOL. 9 No. 3, pp. 65-96 1947

Philips Technical Review

DEALING WITH TECHNICAL PROBLEMS
RELATING TO THE PRODUCTS, PROCESSES AND INVESTIGATIONS OF
N.V. PHILIPS' GLOEILAMPENFABRIEKEN

EDITED BY THE RESEARCH LABORATORY OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN, EINDHOVEN, HOLLAND

PROBLEMS IN PHOTOGRAPHIC REPRODUCTION, IN PARTICULAR OF SOUND-FILMS

by C. J. DIPPEL and K. J. KEUNING.

778.588.3:771.534.553

The resolving power of a film made by the usual photographic methods is limited by the circle of diffusion formed by the grains of the film. Consequently in order to get sharp pictures on the projection screen the images on the film must be of a certain minimum size. If the film carries a sound track and the speed of the projected film is fixed, then, in spite of a so-called cancellation method being applied when recording and copying, this limited resolving power results in a loss in the amplitude of high frequencies. In order to counteract the circle of diffusion it is desirable that the film should have a high gamma, of say 4 or 5. For picture reproduction, however, Goldberg's rule prescribes a gamma in the neighbourhood of 1 or 2. The compromise that has to be reached when a picture film and a sound track have to be copied on a single film by the usual methods makes its influence felt throughout the whole of the present-day technique of cinematography.
A much simpler and less expensive solution of the problem of copying sound-films is offered by a new method of photographic reproduction that was developed in the Philips laboratories during the war. This method is based on the use of a diazonium compound combined with a mercury salt. The most striking features of this method are the extremely high resolving power (1000 lines per millimetre) and the locally variable gamma (between, *e.g.*, 1 and 8). A more detailed description of this method and its possibilities of application will be given in another article to be published in this journal shortly.

During the war a group of research workers in the Philips Laboratories developed an entirely new method of photographic reproduction. Besides other applications to be referred to later on, some of which are quite new, they had in mind the making of copies of sound-film. In one of the next numbers of this journal we hope to give a concise explanation of this new system of reproduction, its characteristic features and the remarkable possibilities of its application. We deemed it advisable to deal first with the main problems arising in photographic reproduction, in particular of sound-films. In fact in the last decades the photographic methods commonly employed — most of them based on the use of silver halide — have reached such a high state of perfection that one may wonder whether it is not presumptuous and unnecessary to advance a new method. We hope, however, that what follows will make it clear that there are indeed drawbacks attaching to the present methods, drawbacks which have led to this new development. From that it will be seen what place it can take in photographic reproduction.

The aspects from which a system of reproduction is to be judged

By speaking of photographic reproduction we have already given expression to the fact that it is not the recording but the copying of sound film that forms the subject of this article. Even though both recording and copying are done photographically there are great differences between the two. The exposure time when recording has to be very short in order to get sufficient definition of the variations of picture and sound, whereas copying can be done, in principle, with any length of exposure. For taking the picture (though it may be in colours) the colour sensitivity of the film material has to be about equal to that of the eye to translate the picture in the right shades of grey. For copying the black and white picture the copying material may have any spectral sensitivity provided it is suitable for the colour of the light source employed, which is likewise primarily arbitrary.

Here we have mentioned some properties that are a matter of indifference in reproduction material.

What are, indeed, of importance are mainly the following points.

In the first place there is the quality of reproduction. A good quality picture must be sufficiently sharp and properly graduated in the half-tones. In sound it is desired to retain the right amplitudes of the high frequencies, which in most systems of reproduction are apt to be sacrificed; there should be no distortion, and the pitch of each note should be absolutely constant (no "whining" as heard with gramophone records).

Another condition for the copying of films is that the cost of the copying material and of the processing must be kept low, for in many cases this is done in mass production; as opposed to the taking of the film, where the cost of the film material is as nothing compared with the cost of the acting, scenery, staging, etc. This requirement will be particularly applicable when it comes to using the combined picture and sound film for other purposes outside the cinema.

Further requirements are that the film copy must be durable and, preferably, not easily liable to damage. Safety film is also to be recommended — hence the preference shown in some cases for the 8 and 16 mm sub-standard films, as these are made exclusively of safety material — though it is not regarded as a condition *sinc qua non.*

Finally we have to mention some properties which need not be specifically demanded of a sound track on a film because they are already inherent in a film. These properties are of importance when comparing the sound recording on a film with that on gramophone records. They are: greater length of uninterrupted playing on a single standard reel; ease of transport and storage of the copies (there is little difference in volume of what is recorded, for the same playing time, on gramophone records and on normal film reels); further the automatic synchronisation of pictures and sound when both are side by side on one film. To fully appreciate what this synchronisation means it is necessary to recall that in the beginning of the history of talking films, when the sound was reproduced from gramophone records, it was very difficult while projecting a film to match the sound exactly to the picture on the screen. Modern developments have solved this problem and synchronisation is now done as part of the copying process; having regard to the quality of the sound it has been found impracticable to record sound and picture simultaneously on the film, so that usually the sound is recorded separately, the picture film and sound track then being copied side by side on one film, taking care that the two are properly synchronised.

The aspects of a system of reproduction brought forward here are not all independent of each other. For instance, we have seen that the sensitivity of the copying material is not of primary importance because one is not essentially bound to making short exposures, but considered from the point of view of the cost of the copying process sensitivity is indeed of some account, because the copying process should not take too long. Of still more importance is the relation between the cost factors and the quality of reproduction, as will be seen from what follows.

The resolving power of the film

The cost of material for a film copy will be all the less according as the size of picture is chosen smaller and the film speed is lower; these are factors allowing of a longer playing time to be compressed into a shorter length of film.

Since, as a rule, it is desired to project a film at the rate of 24 frames per second, the speed of a picture film is fixed by choosing the suitable size of picture (height of frame). In the case of a film carrying only a sound track — as for instance where the Philips-Miller recording system is applied in broadcasting studios [1]) — there is more freedom in the choice of the recording and the projecting speed.

The smaller, however, the frame height and the lower the film speed, the more difficult it becomes to satisfy the requirements for good quality reproduction. This we will explain first in connection with the picture and then in regard to the sound.

The first difficulty encountered in the case of a very small picture is to concentrate on it the high light flux required to produce a sufficient intensity on the screen. Among other requirements is a very high brilliancy of the light source [2]). A second difficulty, however, is one inherent in the film itself, its capability of giving a sufficiently sharp picture on the screen. It is a commonly known fact that it is pointless to go beyond a certain limit in enlarging a photographic picture because if one goes beyond that the enlargement is blurred and no new details are shown. This limit is usually set by the resolving power of the film. To get a very small frame size without sacrificing sharpness of the screen picture the film must possess a very high resolving power.

Since this brings us to a cardinal point in photographic reproduction, it is necessary to give closer consideration to what we mean by resolving power. This may be defined in several ways. The most commonly accepted conception is indicated by the number of lines per millimetre that can be printed on the film by copying a test object without their becoming indistinguishable (*fig. 1*). This, however, is rather a subjective measure, depending also upon various objective factors such as the illumination of the test object, width and contrast of the light and the dark parts, etc. This makes it difficult at times to compare the measures given by various investigators for the resolving power of films. The resolving power of standard positive film as used for making film copies is about 55 lines per mm. That of the Kodak Panatomic film for ordinary photography is somewhat higher, *viz.* 60 lines/mm for normal and 70 lines/mm for so-called fine-grain development; still better are the Kodak "High Contrast Positive" films (100 lines/mm) and "Microfile" (160 lines/mm).

---

[1]) Philips Techn. Rev. 5, 74, 1940.

[2]) See *e.g.* Philips Techn. Rev. 4, 2, 1939 and 8, 72, 1946.

Before dealing with the causes and effects of this limited resolving power we will consider how matters stand also in regard to the sound reproduction. We will confine our attention to variable area sound tracks (*fig. 2*), though similar considerations apply also to variable density tracks. It will easily be realised that broadly speaking the position with regard to sound is analogous to that of the picture, since the recording of sound is again a question of sharp reproduction of certain details. The sound track of a frequency of say 9000 c/s recorded at a film speed of 30 cm/sec. contains 30 sines on 1 mm. To record these 30 sines exactly the film would have to have an exceptionnally high resolving power, much higher than the normal values. If the resolving power is inadequate then, as we shall see farther on, the sines are distorted, so that, firstly, a number of harmonics of the original frequency and — where composite sounds are recorded — also sum and difference frequencies arise (non-linear distortion), and, secondly, the amplitude of the original frequencies is reduced.

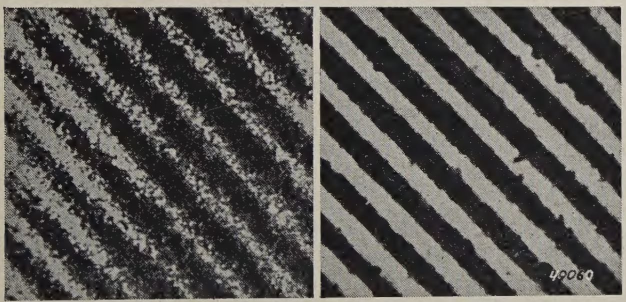
a *b*

Fig. 1. The limited resolving power of a film. When a grating is pictured on the film the lines flow into each other as soon as the distance between the lines on the film becomes too small. *a*) Grating recorded on a film with low resolving power. *b*) The same grating recorded on a film with a much greater resolving power. (Taken from W. Meidinger. Die theoretischen Grundlagen der photographischen Prozesse, J. Springer, Vienna, 1937.)

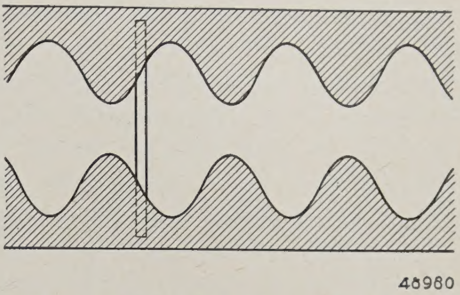

Fig. 2. Variable area sound track recorded on film. This track is played by throwing onto it a narrow beam of light perpendicular to the direction of the film; the light passing through varies with the width of the track and these variations are made audible by means of a photocell and amplifier.

These effects become all the more pronounced according as higher notes are recorded and the speed of the film is reduced. Consequently, for certain admissible values of distortion and of amplitude loss in the high frequencies, the film must have a high resolving power if it is desired to reduce the film speed.

Cause of limited resolving power

Obviously the limited resolving power of a film is related to the grain of the film material. The granular structure of the emulsion [3]) is in itself so fine that its irregularity would still allow of a grating being resolved of much more than 50 or 70 lines per mm. The disturbing effect lies rather in the diffusion of light (also called the circle of confusion) produced when exposing the film and caused by the scattering of light by the grains in the emulsion. Owing to this diffusion there is light playing in a turbid emulsion layer in places where it ought not to be. As a result "image spread" arises to a greater or lesser degree according to the intensity (and also the wavelength) of the light. With the gelatine emulsion commonly used this phenomenon is, it is true, counteracted to a certain extent by the so-called Ross effect, an "image contraction" due to the slight shrinkage of the gelatine where silver is separated. At a certain density [4]) there is a balance between the two effects.

---

[3]) The term "emulsion" is not really correct: we have here a colloidal solution, not of liquid globules, but of solid particles, the AgBr grains in the gelatine. It has, however, become so common to speak of emulsion that we will not depart from it here.

[4]) "Density" is defined as $D = \log I_0/I$, where I is the quantity of the incident light I_0 that the blackened plate allows to pass through.

For practical purposes, however, this density is much too small (*e.g.* about 0.5), The normally required density is absolutely predominated by image spread, so that the adjacent lines of a too fine grating or, in general, details at too short distances apart are caused to blur.

The magnitude of the image spread can be deduced by experimenting with a circular image of about 0.1 mm in diameter reproduced on a film under different exposures E (intensity × exposure time). Such experiments, first carried out by the astronomer **Scheiner**, have shown that with not too small and not too large values of E the diameter d of the photographic picture can be represented by

$$d = a + b \log \frac{E}{E_0} \quad \dots\dots\dots \quad (1)$$

where E_0 is the exposure under which the image spread is just balanced by the **Ross** effect. Since for $E = E_0$ the second term on the right disappears, the constant a equals the true image size that would be obtained in the absence of the phenomenon of image spread and contraction. Hence the image spread s may be written as:

$$s = b \log \frac{E}{E_0}, \quad \dots\dots\dots \quad (2)$$

in which b is an empirically determined constant.

Considering that in recording or in copying a certain density is reckoned with, we express the image spread as a function of the density D. The ratio of D to log E for any photographic negative or positive is indicated by the H & D curve as given in *fig. 3*. This curve is usually characterized by the "gamma", *i.e.* the slope γ of the linear portion. If we define D_0 as the density corresponding to the exposure E_0 (thus where there is no image spread) then:

$$D—D_0 = \gamma (\log E—\log E_0). \quad \dots \quad (3)$$

From (2) and (3) we derive for the image spread [5]).

$$s = \frac{b}{\gamma} (D - D_0). \quad \dots\dots \quad (4)$$

The conclusion of practical importance to be drawn from this is that, other things being equal, a high gamma of the film (of course assuming that b is not appreciably dependent on γ) is favourable for restricting image spread and, consequently, for increasing the resolving power. This is understandable also without formulae. The circle of light produced by diffusion around an exposed spot is strongest in the middle and gets weaker towards the periphery. If a film with high gamma is used (*e.g.* that of the curve in *fig. 3b*) and the prescribed density D in the exposed spot is aimed at, then in the surrounding circle there will only be a perceptible density where the exposure is not much less than in the spot itself; thus the resulting circle of diffusion is limited to a fraction of the aforementioned circle.

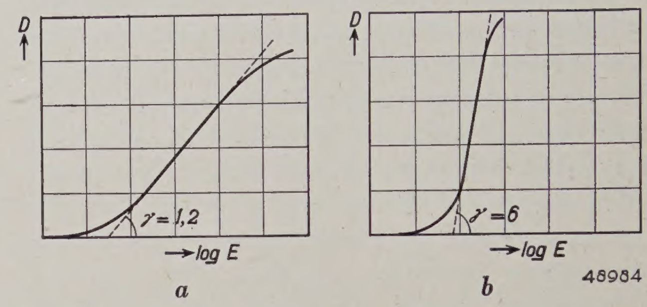

Fig. 3. H & D curve for a photographic negative or positive represented diagrammatically. D = density, E = exposure (= exposure intensity × exposure time), γ = slope of the rectilinear portion of the curve. The gamma depends upon the kind of film and the method of developing. *a*) The case of a low gamma; *b*) the case of a high gamma.

In the case of a sound track (fig. 2), where there is essentially only one density value to be reckoned with [6]), it is indeed only right to aim at the highest possible gamma. For the reproduction of pictures, however, this course cannot be followed, and this we will revert to again later on.

Practical consequences of the limited resolving power

a) *For pictures*

As we have seen, owing to the limitation of the resolving power the size of the picture in a film for projection cannot be chosen below a certain limit. This limit is easily calculated. Reckoning with the resolving power of 55 lines per mm given above for positive films, then details of about 0.02 mm are still sufficiently separated on the film; smaller details run into each other. If the details do run into each other a little it does not matter so long as the cinema patrons do not notice it. Normally the human eye can only discern details and faults in their reproduction when they are viewed under a visual angle of about one minute (at least in the range of brightness with which we are concerned here). This corresponds to 0.1 mm at the distance of clear vision (25 cm), or 6 mm at a distance of 15 metres, which may be taken as the average distance of the patrons from the screen. Therefore the film picture can be projected in a cinema with a linear magnification of 6 : 0.02 = 300 times without the public noticing anything of the consequences of the limited resolving power of the film, except perhaps in the decidedly unfavourable seats close in front of the screen. Reckoning that in a cinema the picture is projected over a width of say 6 metres, the minimum width of the picture on the film will have to be 20 mm. This is in fact approximately the picture width on the 35 mm standard film (the rest is reserved for sound track and perforations).

---

[5]) Equation (3) only holds when the density D_0 (and D) lies on the rectilinear part of the characteristic curve, which it does not as a rule do, because D_0 is too small for that. Still one can then write for s the formula (4), where only the constants b and D_0 (and in eq. (1) also a) are different from those determined experimentally.

[6]) This is not exactly true. The track is recorded by photographing a slit varying in length. Owing to the finite width of this slit one gets on the edge of the track a transitory zone with diminishing density. This implies that in order to avoid distortion it is necessary to apply the **Goldberg** condition, to which we will refer farther on, as a result of which one would no longer be free in the choice of the gamma. In practice, however, image spread appears to be so much greater that the conclusion still holds that one should aim at a high gamma.

Besides the 35 mm film, however, the 16 and 8 mm films have become very popular, with corresponding widths and heights of the picture frames as indicated in *fig. 4.* When projecting at the rate of 24 frames per second the speeds of these three sizes of films 7) are respectively 45.6 cm/sec., 18.3 cm/sec and 9.2 cm/sec.

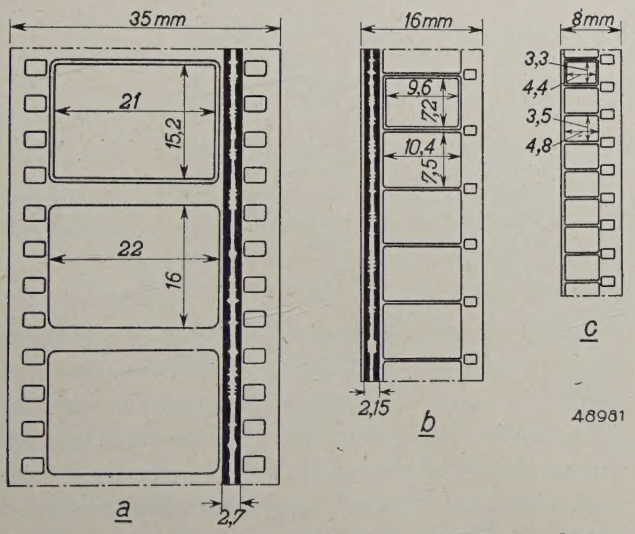

Fig. 4. The three sizes of film commonly used: 35, 16 and 8 mm wide. For each film the width and height of the photographed picture frames are given, the slightly smaller width and height of the film gate in the projector (thus the dimensions of the picture to be projected) and the width of the sound track.

One can see at a glance how much more economical the 16 mm film is than the 35 mm one. For a show of 1 hour 10.5 m^2 of the 16 mm film is required as against 57.5 m^2 of the 35 mm film. If, however, the 16 mm film were to be projected under the same conditions as described above then the limit of resolving power of the positive would undoubtedly already be exceeded and the screen picture would not be absolutely sharp. The fact that nevertheless much use is made of the 16 mm film is due partly to the manner of projection, the public seeing the picture on the screen at a smaller angle (less of the undesired properties but also less of the desired features are seen). It is also in part due to the fact that a certain lack of sharpness in moving pictures need not be decidedly troublesome, for one has, so to speak, no time to notice the lack of sharp definition and the movement itself already distracts attention from the observation of details.

With the 8 mm film, however, even these attenuating circumstances are of no help. Although special film material is used ("reversal" film) with exceptionally fine grain, the picture projected from an 8 mm film is never really sharp. It is therefore practically impossible to reduce the already small frame (4.8 mm) to make room also for a sound track on this size of film. Moreover, the low speed necessary for a film with such a small frame height does not allow of proper sound reproduction. This is why 8 mm films never have sound track and their use is limited to the home kino and to amateurs.

b) *Consequences for sound*

Taking the film speed as fixed, then according to the foregoing the limited resolving power sets a limit to the highest frequencies that can be recorded or copied with sufficient amplitude (here we will disregard a similar limitation due to the finite width of the slit in recording and reproduction).

In sound film technique, in order to avoid this restriction as far as possible, the method of cancellation has been developed. The picture spread occurring in the photographic recording of a sound track causes the black-white limit to be shifted perpendicular to the edge of the sound track. The magnitude of this displacement depends upon the curvature of the original edge. The resulting limitation is therefore by no means congruent with that which would occur without image spread. *Fig. 5b* shows that a sine is distorted to a sugarloaf shape: the peaks are rounded off, the valleys become pointed owing to adjacent delineations

---

7) In the case of the 8 mm film one usually projects at the rate of only 16 frames per second, so that the speed becomes 6.1 cm/sec. In this case 16 frames suffice, because the small pictures are generally projected with a relatively low brightness and the eye is then less sensitive to flickering. The lower speed is not attended with the drawback of sound quality because the usual 8 mm film does not carry any sound track; see below.

flowing into each other. The result is a non-linear distortion of the recorded sound, particularly in the high frequencies. However, when a positive copy is made of the distorted negative sound track there is again image spread, this time causing the track edge to be displaced in the direction opposite to that in the original track. By a suitable choice of density and maybe the gammas of the negative and positive films it is possible to cause the two image spreads to compensate each other fairly well; see *fig. 5c*. The non-linear distortion is then practically eliminated, but, as a comparison of figs. 5*c* and 5*a* shows, there still remains a perceptible loss of amplitude in the high frequencies. Regardless of the application of cancellation, in order to minimise this loss it is an advantage to reduce image spread by choosing a high gamma for the negative and positive films.

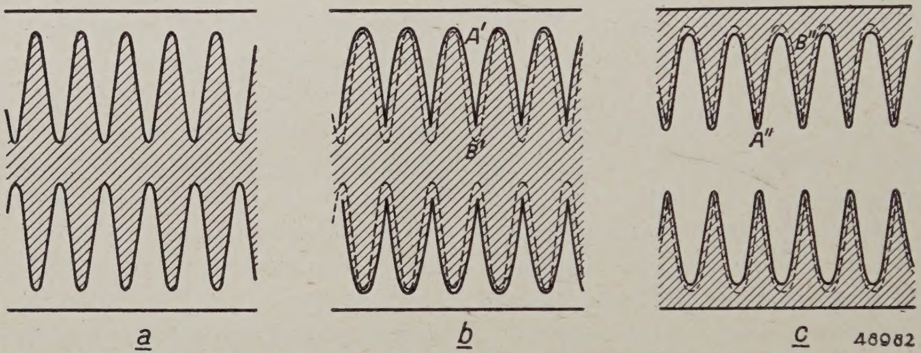

Fig. 5. Cancellation of image spread when recording and copying a sound track, represented diagrammatically.

a) Negative track of a certain frequency as it would be recorded if there were no image spread.

b) Owing to the image spread the inner parts B' of the sines run more or less closer to each other and form sharp peaks, whilst the outer parts A' are rounded off. The ideal shape (*a*) is drawn in dotted lines.

c) When the distorted track (*b*), indicated here by dotted lines, is copied, image spread again occurs, the peaks A'' being widened and the rounded parts B'' narrowed. In this way a practically sinusoidal track can again be produced. It is seen at once, however, that the amplitude of the sine modulation has become smaller than it was in (*a*): the track displacements at B' and B'' are large, those at A' and A'' only very slight.

We will devote for a moment special attention to the problems arising in the copying of sound track recorded by the Philips-Miller system, the principle of which is represented in *fig. 6* [8]).

A celluloid film tape is coated first with a layer of gelatine and over that a thin opaque covering layer. This "Philimil" tape is drawn through underneath a wedge-shaped cutter moving up and down in rhythm with the sound vibrations and thus, by removing the covering layer, cutting in the tape a transparent track of varying width. Here, then, the usual photographic method of recording a sound film is replaced by a mechanical method, with the advantages that the film can be played immediately after its recording (done by optical means, the same as ordinary sound films) and in the reproduction of the high frequencies there is no trouble whatever from the feared image spread. Thanks to the ideally sharp delineation of the track, frequencies up to 8000 c/s can easily be recorded and reproduced, even at a film speed of only 32 cm/sec. Thus one gets a very high quality of sound reproduction. This system has, therefore, already found favour in broadcasting studios for recording commentaries and for other special purposes. For its application in cinemas, however, the "Philimil" recording has to be copied. The same applies if a system for sound reproduction in the home is to be based on the Philips-Miller system, in order to get a step ahead of the gramophone record (on which the high frequencies are reproduced only moderately well). For the mass production of copies only a photographic process can be considered: from the "Philimil" tape, on which the sound track is already positive, a negative has to be printed and from this "intermediate negative" any number of positive copies can be made. If the ordinary photographic methods are applied, however, the phenomenon of image spread is again introduced, and it is just this that has been eliminated in the mechanical recording. In the two photographic copying processes one could apply the principle of cancellation, but then the gain in quality is again partly lost.

---

[8]) For the fundamental principles of this system see Philips Techn. Rev. 1, 107, 1936.

Now in the copying of "Philimil" band a peculiar phenomenon occurs, the so-called lens effect. From *fig. 6* it may be seen that in the recording of the sound track the cutter makes in the transparent gelatine a groove of varying depth and of a triangular cross section. This profile of the gelatine surface acts as a series of lenses which in the copying process break up the homogeneity of the incident light and give rise to alternately very high and very low concentrations of light behind the transparent part of the track. The density differences resulting from this in the negative copy can easily be rendered harmless by seeing to it that in the subsequent copying of the negative also the rather large quantity of light passing through the least dense parts is kept below the threshold of sensitivity of the positive film. The track on the positive will then again be uniformly transparent. The awkward part about this, however, is that in order to get a sufficiently negative density in the parts of the weakest concentration of light one cannot avoid a much more intensive exposure — in practice up to 20 times as much — in the parts with the strongest concentration of light. This results in a marked circle of diffusion and image spread on the track edge where these strong concentrations of light occur.

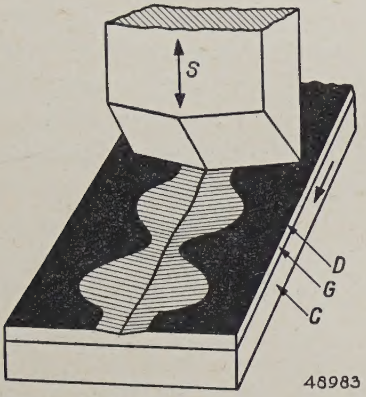

Fig. 6. Principle of the Philips-Miller system of sound recording. The "Philimil" band consists of the celluloid support *C* with a gelatine coating *G* and a very thin opaque covering layer *D*. The cutter *s* vibrating perpendicularly to the band moving in the direction of the arrow cuts out a transparent sound track.

This phenomenon can be counteracted in various ways. For instance, the track could be filled up with a paste or a liquid having the same refractive index as that of the gelatine layer. The easiest way of solving the problem, however, is to choose for the intermediate negative a film with such a high gamma that even with the very intensive light concentrations mentioned above no trouble is experienced from image spread. The desire for a high gamma as already made manifest thus becomes still more imperative. Incidentally, this solution of the problem can only serve if the gamma is not already fixed by the cancellation method first described.

c) *Combination of picture and sound*

In the case of a sound film sound and picture are placed side by side on the same band. In practice this makes it impossible to counteract image spread by aiming at the highest possible gamma, because the gamma is already determined by the picture reproduction. In order to reproduce in the picture by analogous half-tone contrasts all the visual contrasts of the object photographed it is necessary to comply with Goldberg's condition. This condition requires that the product of the gammas of the positive and the negative must be about 1—1.2 [9]). This leads to the compromise of a picture negative with a gamma of abt. 0.8 and a combined sound-picture positive with a gamma of 1.8—2.5, although it would be much more convenient to print the sound with a gamma of 4—6 and to choose for the picture a gamma much nearer to 1. As a matter of fact the choice of gamma also for the picture alone is a compromise, for a different gamma may be desired for the negative film according to the scene taken, and then one would want to adjust the gamma of the copy from scene to scene. Since the gamma of a certain photographic film can only be influenced to any appreciable extent by the developing times, and this can hardly be varied every time for different parts of the film, one must be satisfied with the choice of one particular mean gamma.

Reviewing the situation as a whole we may say that a high resolving power of the film is desirable both for the picture and for the sound track; that a high gamma is favourable but that in the case of a combination of picture and sound the Goldberg condition makes it necessary to arrive at a compromise in this respect. This compromise makes its influence felt throughout the whole of the present-day technique of cinematography. The new system of reproduction referred to in the beginning of this article offers a better solution to this problem.

The new reproduction system

This system is based not on an emulsion of AgBr or AgCl in gelatine but on a combination of a light-sensitive diazonium compound with a mercury salt. The film base is saturated in a solution of this mixture. The system can be applied with different bases. A thin band of cellophane, for instance, is highly suitable for the purpose. After exposure the latent picture obtained is developed by so-called physical development, metallic silver from a solution of a silver salt being deposited on the exposed parts (in contrast to the normal, so-called chemical developing where a silver compound present in the film is reduced to metallic silver in the exposed parts). In this way one obtains a perfectly durable negative image consisting of silver. Compared with the usual photographic negative material, the sensitivity of the new material is low. There can therefore be no question of its competing with silver-bromide emulsion for the taking of films, also because of its sensitivity being restricted to ultra-violet light. For making copies, on the other hand, this mercury diazonium system possesses excellent properties. Since there is here no question of an emulsion but of a homogeneous solution, there are no grains and the film has a resolving power of 1000 lines per mm [10]. The gamma for the dry material is exceptionally high, *viz.* 6—8. A most remarkable feature, however, is that it can easily be adjusted within a wide range already during the exposure, it being possible to reduce it to a very low level.

---

[9]) Regarding the Goldberg condition see, for instance, Philips Techn. Rev. **5**, 51, 1940. Roughly speaking, this condition expresses nothing more than what is known to any photographer, that a "hard" positive paper must be used for printing from a "soft" negative, and *vice versa*.

Thanks to this very high resolving power and the possibility of having a very high gamma, the system is ideal for the reproduction of sound track and also for the photographic multiplication of printed matter and suchlike, where the extremely low cost of the base and of the sensitized material may prove to be of importance. The specific problems arising in the copying of Philips-Miller film are solved by the new system in the simplest possible way. The fact that, if desired, the gamma can be reduced to a low value makes it possible to get good reproductions also of half-tone pictures. Furthermore, the adjustability of the gamma during exposure creates the entirely new possibility of printing picture and sound on the same band with a different gamma and then developing both together.

This brief account of the subject will presumably give rise to more questions than it answers. Our intention here has been only to arouse the reader's interest. Further particulars will be found in one of the next issues of this journal, where we intend to go more closely into the new system of photographic reproduction.

---

[10] The question whether the film is capable of resolving still finer gratings could not be determined because the optical system used for the necessarily reduced recordings cannot itself resolve more than 1000 lines per mm.

ELECTROMAGNETIC CAVITY RESONATORS

by G. de VRIES.

538.565

The forms of oscillation of certain electromagnetic flat cavity resonators are discussed, *i.e.* resonators which may be considered as two-dimensional; namely the forms of oscillation of square plane cavity resonators and the non-rotation-symmetrical forms of oscillation of round plane cavity resonators. Further, the forms of oscillation of three-dimensional cavity resonators are dealt with. The case is then discussed of two coupled cavity resonators, the significanee of which in high-frequency technology is analogous to that of coupled oscillation circuits in the region of lower frequencies, namely to that of a band filter.

As has already been explained several times in this periodical [1][2][3]), ordinary oscillation circuits consisting of concentrated self-inductions and capacities are not suitable for the region of very short waves. The quality factor and the resonance resistance of such oscillation circuits are much too small in that region. Recourse is therefore had to the use of other electrical resonators not having these objections. Very important among these resonators are Lecher systems, *i.e.* two parallel conductors close together or concentric, and cavity resonators, *i.e.* empty spaces surrounded by metal walls. Lecher systems have already been discussed in detail in this periodical [2]). Recently a detailed explanation was also given of the properties of a certain group of cavity resonators [3]), namely of the cavity resonators which are in the form of solids of revolution and in which, moreover, the dimension in the direction of the axis of revolution is small compared with the dimensions perpendicular to it (*fig. 1a*). We have called this kind of cavity resonators round flat cavity resonators. We were then concerned exclusively with the rotation-symmetrical oscillations in which the current and the voltage depend only on the distance to the axis of revolution.

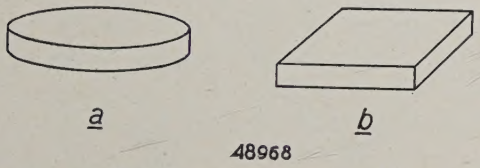

Fig. 1. *a*) A round plane cavity resonator, *b*) a square plane cavity resonator.

In this article we shall go somewhat farther. We shall begin with the consideration of square cavity resonators and especially of such which are not only flat, *i.e.* thin, but also plane, *i.e.* have everywhere the same thickness (fig. 1*b*). The forms of oscillation of square plane cavity resonators are naturally no longer rotation-symmetrical. With our knowledge of square plane cavity resonators we shall then be able to ascertain the characteristics of the non-rotation symmetrical forms of oscillation of the round plane cavity resonators. Moreover, this knowledge will make it possible for us to find the form of oscillation of a cube by an obvious generalization. Thus beginning with a practical two-dimensional case (square plane cavity resonator) we can arrive at conclusions about a three-dimensional case (cube). After a few remarks about the quality factor and the resonance resistance of cavity resonators we shall in conclusion discuss coupled cavity resonators.

Forms of oscillation of square plane cavity resonators

Qualitative considerations

We shall first mention briefly the result previously obtained (see footnote [3]) for the distribution of current and voltage in the rotation-symmetrical forms of oscillation of a round plane cavity resonator. Such a cavity resonator has a "bottom" and a "cover". We shall consider the voltage between a point of the cover and the point on the bottom lying directly beneath it. The current at two such points is equal in value and opposite in direction. It is thus sufficient to speak about the current, for instance, in the cover, as we shall do here.

The centre of the cover is a point of high voltage [4]); at the outer edge the voltage is zero. The current, on the other hand, is zero at the centre; at all other points it is radially directed and reaches a maximum at the outer edge. The current density is a maximum at a short distance from the edge and then decreases slightly towards the edge, but not very strongly, so that the total current, which is equal to the product of current density and circumference, continues to increase from that point to the edge. This holds for all rotation-symmetrical oscillations which further differ from each other in the number of current maxima and zero points along the radius between the centre of the cavity resonator and its outer edge. In the case of the fundamental oscillation the current increases from the centre towards the edge. In *fig. 2a* several current lines are drawn for that case, whereby the width of the lines has been chosen proportional to the current density. In fig. 2*b* the distribution of the current density is shown for the lowest but one rotation-symmetrical oscillation.

---

[1]) C. G. A. von Lindern and G. de Vries, Resonance circuits for very high frequencies, Philips Techn. Rev. **6**, 217, 1941.

[2]) C. G. A. von Lindern and G. de Vries, Lecher Systems, Philips Techn. Rev. **6**, 241, 1941.

[3]) C. G. A. von Lindern and G. de Vries, Flat cavity resonators as electrical resonators, Philips Techn. Rev. **8**, 149, 1946.

[4]) In the previous article (see footnote [3])) it was mainly a question of flat cavity resonators with a circular hole in the centre of cover and bottom. Here, for the present, we are discussing cavity resonators with no hole.

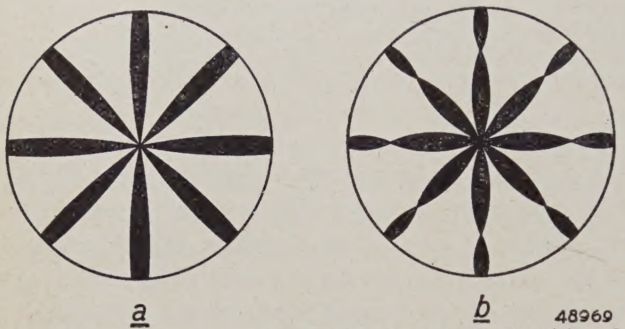

Fig. 2. Representation of the fundamental oscillation (*a*) and of a higher characteristic oscillation (*b*) of a round plane cavity resonator. A few current lines have been drawn whose width is proportional to the current density. Due to the rotational symmetry all the "lines" are congruent.

When we now attempt to sketch a current density distribution for the fundamental oscillation of a square plane cavity resonator it may be assumed that near the centre of the cover the situation will not be very different from that in the case of round plane cavity resonators. The form of the current lines near the edge is determined by the condition that the current lines must be perpendicular to the edge. This condition, whose derivation we shall not give here, enables us to draw the current distribution of a square plane cavity resonator in the neighbourhood of the edge. In this way a picture of the current distribution is obtained which agrees very well with that found by calculation (see *fig. 3*).

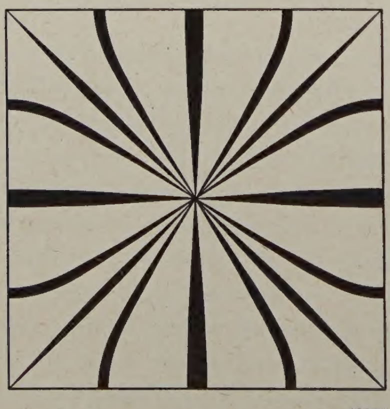

Fig. 3. Representation of the fundamental oscillation of a square plane cavity resonator. Only a few current lines have been drawn whose width is proportional to the current density. The latter differs everywhere from zero except at the centre and at the four corners of the square.

Since over a large part of the square cavity resonator the current distribution does not differ from that of a round cavity resonator, the characteristic frequency will also not deviate too much. In the main the difference between the two kinds of cavity resonator amounts to the fact that in the case of the square plane resonator "corners" have been added. Now close to the edge the voltage is low, so that it is not so much the "capacity" as the "self-induction" of the corners that is decisive for the change in the value of the characteristic frequency. In as far as it is possible to define a conception such as the "average self-induction" of the cavity resonator, such a quantity would become larger by the addition of the corners and the wavelength would thus also become larger. The mathematical theory, which will be discussed in the following, confirms this: while the wavelength of the fundamental frequency of a round cavity resonator is equal to $2{\cdot}61a$ (a is the radius of the round cavity resonator), that of the square one is equal to $2a\sqrt{2} = 2{\cdot}82a$ ($2a$ is the side of the square cavity resonator).

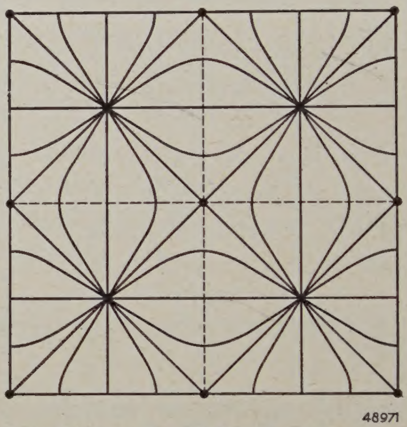

Fig. 4. The occurrence of a higher characteristic oscillation of a square plane cavity resonator. This may be conceived of as the result of placing four smaller square resonators side by side, each of which executes the fundamental oscillation. The points indicate the spots where the current density becomes equal to zero.

Now that we have the figure of the current distribution corresponding to the fundamental frequency of a square plane cavity resonator we can go farther. Let us, for example, place four square plane cavity resonators (side $2a$) side by side in such a way that they form a large square (side $4a = 2b$); see *fig. 4*.

If we then arrange matters in such a way, for example, that the voltage at the centre of the cover of the upper left-hand resonator is positive and that of the upper right-hand resonator negative, the lower right hand positive and the lower left-hand negative, and that these voltages are equal in absolute value, the adjacent currents at a partition are equal and in the same direction. We may therefore omit the partitions without changing the current distribution of the four cavity resonators. From this we conclude that the current distribution in fig. 4 corresponds to a higher form of oscillation of the large square cavity resonator. The longest wavelength for this cavity resonator would be equal to $2b\sqrt{2}$; here, however, we are concerned with a form of oscillation to which the wavelength $2a\sqrt{2} = b\sqrt{2}$ corresponds.

By again placing side by side four square cavity resonators, each of which is in the higher oscillation state just described, we would in a similar manner find the current distribution corresponding to a wavelength four times as small as the wavelength λ_0 of the fundamental oscillation.

It is clear that in this manner it would be possible to find the current distribution for a whole series of forms of oscillation whose wavelength λ would be given by the formula $\lambda = \lambda_0/2^n$ ($n = 0, 1, 2, \ldots$). It must not, however, be thought that all the forms of oscillation of a square plane cavity resonator would then have been found; from the mathematical treatment, to which we shall now pass on, it will be found that many other forms of oscillation are still possible.

Mathematical treatment

The strict theory of electromagnetic cavity resonators is based directly on Maxwell's equations. It is often possible, however, to deduce correct results in a more elementary way, especially for plane cavity resonators.

Thus, for example, in the article already referred to (see footnote [3]) we have developed the mathematical theory of round flat (not only plane!) cavity resonators beginning with the ordinary "cable equations" applying to a Lecher system consisting of two parallel conductors (see *fig. 5a*):

$$\frac{\partial i}{\partial x} = -C\frac{\partial V}{\partial t}, \qquad \frac{\partial V}{\partial x} = -L\frac{\partial i}{\partial t}; \quad . \quad . \quad (1)$$

in these equations C and L are the capacity and the self-induction, respectively, per unit of length; V is the voltage between a point of the upper conductor and the point of the lower conductor directly below it; i is the current in the upper conductor; that in the lower conductor is then $-i$; finally x is the coordinate in the direction of the conductors and t is the time. By rotating such a Lecher system a round plane cavity resonator (see fig. 5*a*) is obtained, and by rotating a Lecher system of a general form, such as that of fig. 5*b*, a round flat cavity resonator is obtained which is no longer plane and which, moreover, has a hole at the centre. It is thus understandable that the rotation-symmetrical modes of oscillation of round flat cavity resonators are also described by equation (1), on the understanding that C and L are, respectively, the capacity and self-induction per ring of 1 cm width, and that in general it will be functions of the radius r which here take over the rôle of the coordinate of length x.

In the case of a square plane cavity resonator the method of treatment sketched above could also be used. We shall indicate briefly how this should be done, although it is not of great importance since the rigorous calculation directly from Maxwell's equations is no more laborious. Again we introduce the voltage V between a point on the "cover" and the point on the "bottom" directly beneath it. Further we introduce the current density which at every point, for instance, of the cover of the square resonator has an x and a y component, i_x and i_y; we assume at the same time that the coordinate axes are parallel to the sides of the square. The current density components in the bottom will then be $-i_x$ and $-i_y$. Finally we introduce the capacity C and the self-induction L per cm² of base plane. It is then obvious, analogous to equation (1), to write the following equations:

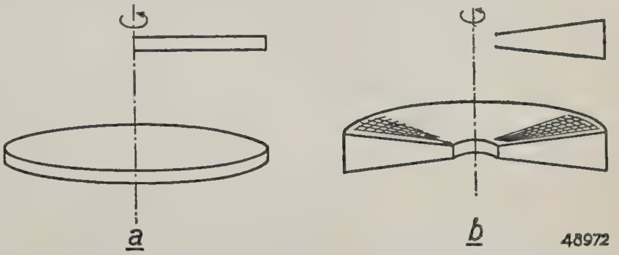

Fig. 5. The formation of a round flat cavity resonator by the rotation of a Lecher system short-circuited at one end: *a*) plane cavity resonator, *b*) resonator of a more general form.

$$\left.\begin{aligned} \frac{\partial V}{\partial x} &= -L\frac{\partial i_x}{\partial t}, \\ \frac{\partial V}{\partial y} &= -L\frac{\partial i_y}{\partial t}, \\ \frac{\partial i_x}{\partial x} + \frac{\partial i_y}{\partial y} &= -C\frac{\partial V}{\partial t}. \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (2)$$

The value to be substituted here for C is obvious: if h is the thickness of the cavity resonator then C is the capacity of a condenser of 1 cm² area and with a plate distance h. On the other hand it is difficult to assign to L a similar physical significance.

Further consideration shows, however, that L follows from the relation $LC = 1/c^2$, in which c is the velocity of light.

The conditions at the edge to be satisfied by the solutions of (2) have already been mentioned: at all points on the edge V must equal 0 and the current density must there be perpendicular to the edge. When the side of the square cavity resonator has the length $2a$ and the origin of the x, y coordinates is situated at the centre of the square, the second condition means that $i_x = 0$ for $y = +a$ and $i_y = 0$ for $x = \pm a$.

We shall not discuss here all the periodic solutions of (10). As an example we give the following family of solutions:

$$\left.\begin{aligned} V &= V_0 \cos k_1x \cdot \cos k_2y \cdot \cos \omega t, \\ i_x &= -V_0 \frac{k_1}{\omega L} \sin k_1x \cdot \cos k_2y \cdot \sin \omega t, \\ i_y &= -V_0 \frac{k_2}{\omega L} \cos k_1x \cdot \sin k_2y \cdot \sin \omega t, \end{aligned}\right\} \quad (3)$$

with

$$k_1^2 + k_2^2 = \frac{\omega^2}{c^2}; \quad . \; . \; . \; . \; . \; . \; . \; . \; . \quad (4)$$

2ω is here the frequency of the oscillations and V_0 an arbitrary constant. In order to satisfy the boundary conditions the following must hold:

$$a\,k_1 = m\,\frac{\pi}{2}, \quad a\,k_2 = n\,\frac{\pi}{2} \quad (m,\; n = 1,\; 3,\; 5\; \ldots) \quad (5)$$

The mode of oscillation sketched in fig. 3 corresponds to $m = n = 1$. Since the wavelength λ is equal to $2\pi c/\omega$, with the help of (4) and (5) it is easy to confirm the statement we made about this mode of oscillation, namely that in this case $\lambda = 2a\sqrt{2} = 2{\cdot}82\; a$.

On the other hand the mode of oscillation represented in fig. 4 is an example of a solution *not* given by a formula of the form (3). This mode of oscillaton corresponds to the following family of solutions:

$$\left.\begin{aligned} V &= V_0 \sin k_1x \cdot \sin k_2y \cdot \cos \omega t, \\ i_x &= -\frac{V_0}{\omega L} \cos k_1x \cdot \sin k_2y \cdot \sin \omega t, \\ i_y &= -\frac{V_0}{\omega L} \sin k_1x \cdot \cos k_2y \cdot \sin \omega t, \end{aligned}\right\} . \quad (6)$$

with

$$k_1^2 + k_2^2 = \frac{\omega^2}{c^2}$$

and

$$ak_1 = m\pi, \quad ak_2 = n\pi \quad (m,n = 0, 1, 2\; \ldots).$$

For $m = n = 1$ we obtain the mode of oscillation in question. The wavelength is now $\lambda = a\sqrt{2} = 1{\cdot}41\; a$.

Non-rotation-symmetrical form of oscillation of a round plane cavity resonator

The mode of oscillation of a square cavity resonator represented in fig. 4 could not be derived directly from a rotation-symmetrical form of oscillation of a round cavity resonator. Conversely, however, from this higher mode of oscillation of a square resonator we may now derive a non-rotation-symmetrical form of oscillation of a round plane cavity resonator. It is not difficult to make a sketch on the basis of fig. 4 of the current distribution in a round resonator for the mode of oscillation in question. In the main this current distribution will present the same picture as in the case of the square resonator, but near the edge we must alter the current lines somewhat so that they are perpendicular to it and thus radial in direction. This leads to *fig. 6b*. The application of a more rigorous mathematical theory fully confirms the correctness of this figure.

It is useless to attempt to force the theory for the non-rotation-symmetrical forms of oscillation of round resonators into the mould of the cable equations. The self-induction per cm² was already a rather artificial conception; it would be still more unnatural to speak of the self-induction of a surface element $r \cdot dr \cdot d\varphi$. We shall thus give directly the results produced by the theory based on Maxwell's equations. We introduce the polar coordinates r and φ on the plane of the cover of the round cavity resonator. The voltage V is then given by

$$V = V_0\, J_m\,(kr) \cos m\varphi \cdot \cos \omega t,\; (m = 0, 1, 2)\; . \quad (7)$$

where $J_m(x)$ represents the Bessel function of the order m and

$$k = \frac{\omega}{c} = \frac{2\pi}{\lambda};$$

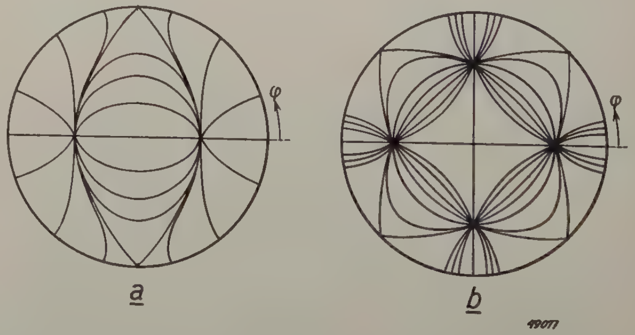

Fig. 6. Diagram of a non-rotation-symmetrical form of oscillation of a round plane cavity resonator. Only a few current lines are drawn here. *a*) The mode of oscillation where the current density near the edge of the resonator is proportional to cos φ. *b*) The mode of oscillation where the current density near the edge of the resonator is proportional to cos 2φ.

λ and ω are the wavelength and the angular frequency respectively.

At the edge of the cavity resonator the voltage must become equal to zero. When the radius of the resonator is R the following condition must be satisfied:

$$J_m(kR) = 0 \quad (m = 0, 1, 2 \ldots). \quad . \quad . \quad . \qquad (8)$$

Since the Bessel function $J_m(x)$ has an infinite number of roots for every value of m, with a given value of m we obtain an infinite number of values of k satisfying equation (8). With a given value of m we may arrange these roots according to increasing size and assign an order number p (1, 2, 3, etc.) to each root. Since m itself may also have an infinite number of values, we thus obtain a doubly infinite series of characteristic oscillations. Values of $kR = 2\pi R/\lambda$ which correspond to several of the lowest characteristic frequencies, *i.e.* to several of the smallest values of m and p, are given in *table I*. For $m = 0$ the voltage distribution is rotation-symmetrical, for $m = 1$ the voltage is proportional to $\cos \varphi$, for $m = 2$ proportional to $\cos 2\varphi$, etc.

Table I

Several values of $2\pi R/\lambda$ for which $J_m(2\pi R/\lambda = 0$; m is the order of the Bessel function J_m, p the order of the zero point of J_m.

m \ p	1	2	3	4
0	2.405	5.520	8.654	11.792
1	3.832	7.016	10.173	13.323
2	5.135	8.417	11.620	14.796

The formulae for the components i_r and i_φ of the current density are as follows:

$$i_r = \frac{V_0}{h}\frac{c}{4\pi} J_m{}'(kr) \cos m\varphi \cdot \sin \omega t, \quad . \qquad (9a)$$

$$i_\varphi = \frac{V_0}{h}\frac{c}{4\pi}\frac{J_m(kr)}{kr} m \cdot \sin m\varphi \cdot \sin \omega t; \qquad (9b)$$

where h is the thickness of the cavity resonator and c the velocity of light (if i and V_0 are expressed in amperes and volts respectively, the lengths in cm, then $1/c = 30$ ohms); the derivative of J_m is $J_m{}'$.

The current distribution in fig. 6*b* corresponds to $m = 2$ and $p = 1$ (*c.f.* table I). It may therefore be seen that this form of oscillation is not the first following the rotation-symmetrical fundamental oscillation ($m = 0, p = 1$). The current distribution corresponding to $m = 1$, $p = 1$ is represented in fig. 6*a*. The forms of oscillation for which $m = 0$ and $p = 1, 2, 3, \ldots$ are all rotation-symmetrical and have already been discussed in detail in a previous article (see footnote [3]).

Because of (8) $i_\varphi = 0$ at the edge of the cavity resonator, *i.e.* the current density is radially directed in agreement with the condition formulated in the treatment of square cavity resonators. From (9a) it then follows that the current density close to the edge is proportional to $\cos m\varphi$. If inside the cavity resonator we introduce along its edge a number of loops whose plane contains the axis of revolution of the cavity resonator, currents will be induced in those loops whose intensity is also proportional to $\cos m\varphi$. This can easily be demonstrated by connecting small lamps to such loops. For the rotation-symmetrical form of oscillation all the lamps will burn (*fig. 7, left*); for the $\cos \varphi$ mode of oscillation two groups will burn: at $\varphi = 0$ and at $\varphi = \pi$ (fig. 7 *middle*); for the $\cos 2\varphi$ oscillation four groups will burn: at $\varphi = 0$, $\varphi = \pi/2$, $\varphi = \pi$ and $\varphi = 3\pi/2$ (fig. 7, *right*). The form of oscillation prevailing depends upon the frequency applied to the cavity resonator.

With the above the most important facts about the forms of oscillation of square and round flat cavity resonators have been stated. The practical significance of the higher forms of oscillation of cavity resonators in general will be explained presently with reference to an example. First, however, we shall examine the appearance of the forms of oscillation in the case of "three-dimensional" cavity resonators, *i.e.* cavity resonators which can no longer be considered flat.

Three-dimensional cavity resonators

We shall begin with cavity resonators in the form of a cube. We have seen (see footnote [3]) that the oscillations of a Lecher system of length a and open at one end could be described by the formulae

$$\left.\begin{aligned} V &= V_0 \cos kx \cos \omega t, \\ i &= i_0 \sin kx \sin \omega t, \\ k &= \omega/c, \\ ak &= m\frac{\pi}{2} \quad (m = 1, 3, 5, \ldots). \end{aligned}\right\} \quad . \quad . \qquad (10)$$

For a square flat cavity resonator we found above (see equations (3) and (4)) formulae of the form

$$\left.\begin{aligned} V &= V_0 \cos k_1 x \cos k_2 y \cos \omega t, \\ i_x &= i_{0_x} \sin k_1 x \cos k_2 y \sin \omega t, \\ i_y &= i_{0_y} \cos k_1 x \sin k_2 y \sin \omega t, \\ k_1{}^2 + k_2{}^2 &= \omega^2/c^2, \\ ak_1 = m\frac{\pi}{2},\ ak_2 &= n\frac{\pi}{2} \quad (m, n = 1, 3, \ldots), \end{aligned}\right\} \qquad (11)$$

where $2a$ is the length of the side of the square,

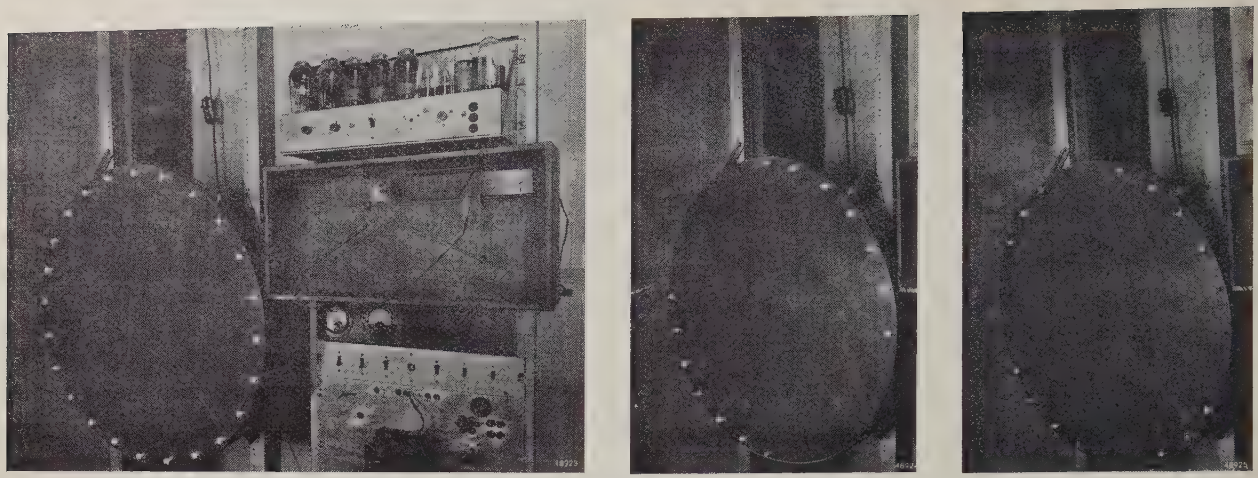

Fig. 7. *Left:* A round plane cavity resonator which executes the fundamental oscillation (*c.f.* fig. 2). This form of oscillation is rotation-symmetrical, as is demonstrated by the fact that the lamps along the edge all burn at the same intensity. These lamps are connected with loops inside the cavity resonator. The cavity resonator is excited with the help of a loop (it projects from the slot in the side wall of the resonator), which is connected with an oscillator (on the right in the figure).

Centre: A round plane cavity resonator in a non-rotation-symmetrical state of oscillation in which the current density near the edge is proportional to cos φ (*cf.* fig. 6*a*). Only two groups of lamps burn: those at $\varphi = 0$ and those at $\varphi = \pi$. The oscillation takes place in such a way that the loop exciting it is situated at a position of maximum current density on the edge.

Right: A round plane cavity resonator in a state of oscillation where the current density near the edge is proportional to cos 2φ (cf. fig. 6*b*). Four groups of lamps are lighted, those at $\varphi = 0$, $\varphi = \pi/2$, $\varphi = \pi$ and $\varphi = 3\pi/2$.

This was the case for "one-dimensional" and "two-dimensional" systems. It is now reasonable to assume that for a cube, whose sides are also squares, analogous formulae will be applicable. We thus assign to each wall a current density distribution corresponding to equation (11). If the x, y, z axes are parallel to the edges (length $2a$) of the cube and if the origin of the coordinates is at the centre of the cube, the following formulae should apply (we omit the factor sin ωt): for the walls which are parallel to the x, y plane (the x, y walls)

$$\left.\begin{array}{l}
i_x = i_1 \sin k_1x \cos k_2y, \\
i_y = i_2 \cos k_1x \sin k_2y; \\
\text{for } y,\ z \text{ walls} \\
i_y = i_3 \sin k_2y \cos k_3z, \\
i_z = i_4 \cos k_2y \sin k_3z; \\
\text{for } z,\ x \text{ walls} \\
i_z = i_5 \sin k_3z \cos k_1x, \\
i_x = i_6 \cos k_3z \sin k_1x; \\
\text{and the following equations must} \\
\text{thereby be satisfied:} \\
k_1^2 + k_2^2 + k_3^2 = \omega^2/c^2 \\
\text{and} \\
ak_1 = m\pi/2,\ ak_2 = n\pi/2,\ ak_3 = p\pi/2 \\
\qquad (m,n,p = 1,3,5\ldots)
\end{array}\right\} \quad \ldots \quad (12)$$

The current density has the same distribution, except for the sign, in two parallel walls of the cube as in the case of a flat square cavity resonator.

Are the formulae (12) now indeed correct? Before we answer this question we would call attention to the fact that among the equations (12) we have given no formula for the voltage V. This was not done without reason. While the current density in the walls of the cavity resonator remains a physical quantity in our somewhat daring generalization from two to three dimensions, it is difficult to picture directly what is to be understood by voltage in the case of a cubic cavity resonator. This brings us to a fundamental question, namely to what extent it is possible to speak of a voltage, *i.e.* of a potential difference, outside the field of electrostatics and direct currents. We must consider that question for a moment.

The voltage V, *i.e.* the potential difference between two points (for instance A and B), can be defined unambiguously in electrostatics due to the fact that the value of the integral $\int_B^A E_s\, d_s$, in which E_s is the component of the electrical field tangential to the path of integration, is independent of the path of integration. The voltage between A and B is then by definition

$$V_{AB} = \int_A^B E_s \,\mathrm{d}s$$

along any given path. This definition can still be used in the same form outside the field of electrostatics as long as one is only concerned with direct currents. As soon, however, as alternating currents are considered certain new conventions must be adopted in order to be able to continue to use the concept of "voltage".

Let us consider for example the simple case of a conductor in the form represented by *fig. 8*; for the sake of simplicity we shall disregard the resistance of the conductor. When an alternating current flows in the conductor the value of $\int_A^B E_s \, ds$ will certainly be different along the various dotted paths. According to the induction law the integral $\int E_s$ ds along a closed path — for instance from B to A along one path and back from A to B along another — is proportional to the derivative with respect to time of the magnetic flux through the area enclosed by this path. And this derivative with respect to time is certainly not equal to zero, since the magnetic field of an alternating current changes with the time. The "voltage" between A and B measured along different paths should then actually be different. Nevertheless, one may usually still continue to speak of voltage, if it is agreed that one means the integral $\int_A^B E_s$ ds along the shortest path between A and B. This agreement is not merely formal, but usually also practically justified, because it need not by any means be exactly the shortest path. As long as the derivative with respect to time of the magnetic flux enclosed by the shortest path from A to B (dotted line I in fig. 8) and a slightly deviating path (for instance dotted line II or III) is small compared with $\int_A^B E_s$ ds along the shortest path, the difference between the voltages measured along these two paths will remain relatively small. One will then indeed continue to speak of the voltage. If the wavelength is approximately equal to the distance AB or smaller, the magnetic field strength will in general change very much for distances of this order of magnitude. As a result the difference between the voltage measured along the shortest path from A to B and the voltage measured for instance along the dotted line II may be proportionately large. But in many cases (by no means in all), also in the region of very high frequencies it is possible to speak of the voltage, in the sense of our convention, when systems are considered whose dimensions are much smaller in one direction than in the other.

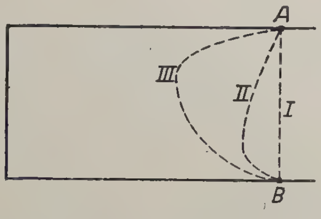

Fig. 8. The voltage between two points A and B of a conductor in which an alternating current flows is not unambiguously defined: the voltage measured along different paths (dotted lines I, II and III) is different.

Such a case is now met with in the case of flat cavity resonators. Since the thickness of a flat cavity resonator is much smaller than the wavelength which corresponds to not all too high characteristic frequencies, the result is that the value of $\int_A^B E_s$ ds along the shortest path between the "bottom" and the "cover" differs relatively little from the value of the integral along a slightly different path. Therefore it was possible to speak of the voltage between a point of the "cover" of a flat cavity resonator and the point in the "bottom" directly under it. Because, however, in the case of a cubic cavity resonator the distance between two such points is not small compared with the other dimensions, it is hardly reasonable to speak of a voltage in that case [5]). This is true, of course, not only for a cubic cavity resonator but also for other "three-dimensional" cavity resonators.

The oscillations must then be described with the help of the electrical and magnetic field strength. The current density in the walls of the cavity resonator remains a reasonable idea: it has the significance of the current density induced by the variable magnetic field. The above formulae (12) actually correspond to certain forms of vibration of the cubic cavity resonator and give us an idea of the distribution of current density in its walls. It is found that there are relations between the constants i_1, i_2, $i \ldots i_6$, which we shall not go into here. In *fig. 9* the distribution of the current density is sketched for two modes of oscillation both of which correspond to $m = n = p = 1$ in equation (12), but to different values of the i's.

For cylindrical cavity resonators with circular cross section formulae are obtained for the current density in the walls which are again a generalization of equation (9) for round flat cavity resonators. Further at the cost of much calculation the current distribution can be studied for spheres, ellipsoids and elliptical cylinders, without enriching the qualitative picture already formed of the possible modes of oscillation of a cube. Rather than enter into these mathematically fairly difficult questions, we shall at the end of this article tell something about a case which in a different respect is more complicated than those considered until now and which does indeed widen our field of vision somewhat.

---

[5]) For certain forms of oscillation of a cube, namely when the electric field is everywhere perpendicular to a side wall of the cube, the voltage could be defined in a fairly natural manner as the integral $\int E_s$ ds along the line of force. A slight deviation from this path of integration would then, however, result in an appreciably different value of the integral, so that this definition would only have a formal significance.

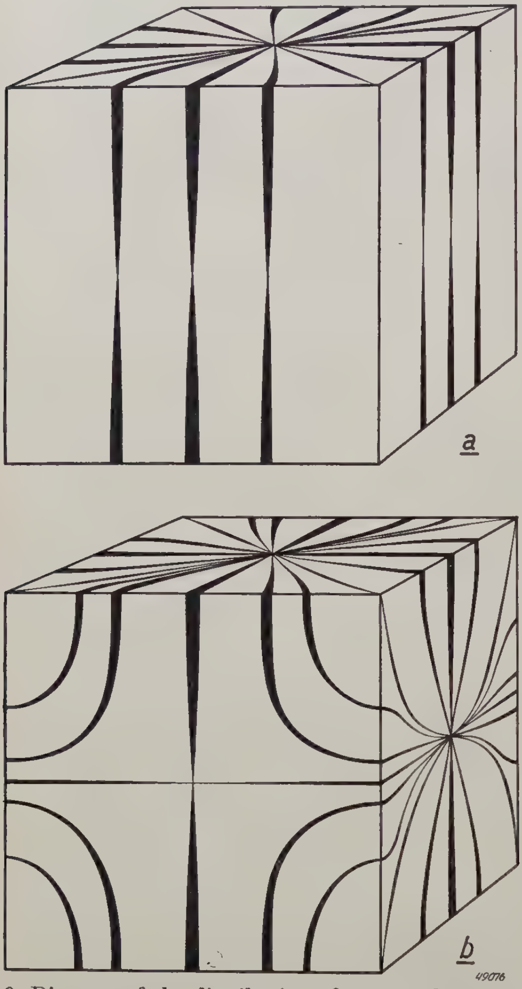

Fig. 9. Diagram of the distribution of current density in the walls of a cubic cavity resonator. The width of the "lines" is drawn proportional to the current density. The forms of oscillation *a*) and *b*) both correspond to $m = n = p = 1$ in equation (12); the amplitudes $(i_1, \dots i_6)$, however, are different in the two cases.

The practical significance of the higher forms of oscillation of the cavity resonators is illustrated by *fig. 10.* In this figure a so-called induction-tube oscillator may be seen, the principle of which has already been discussed in a previous article (see footnote [3]). The cavity resonator here has the function of an oscillator circuit in which the oscillations are excited by the induction tube mentioned. The latter is situated in a hole in the cavity resonator at the position of a voltage maximum. If it is now desired to obtain an oscillator with a high power, two induction tubes could be placed side by side in the hole. Such an arrangement, however, meets with all kinds of objections. Use is therefore made of the fact that with higher forms of oscillation a cavity resonator has more than one voltage maximum. Holes are made at the positions of these maxima and each induction tube is placed in a separate hole.

We shall now discuss the quality factor and the resonance resistance of cavity resonators in general. In practice these two quantities often determine which cavity resonators are to be considered for a given technical purpose. Whether one or another cavity resonator is chosen among those having a suitable quality factor and resonance resistance usually depends chiefly upon the structural requirements which the apparatus involves.

Quality factor and resonance resistance of cavity resonators

The quality factor of a cavity resonator can be considered as a measure of the sharpness of resonance, *i.e.* of the selectivity of a cavity resonator. The larger the quality factor the greater the selectivity.

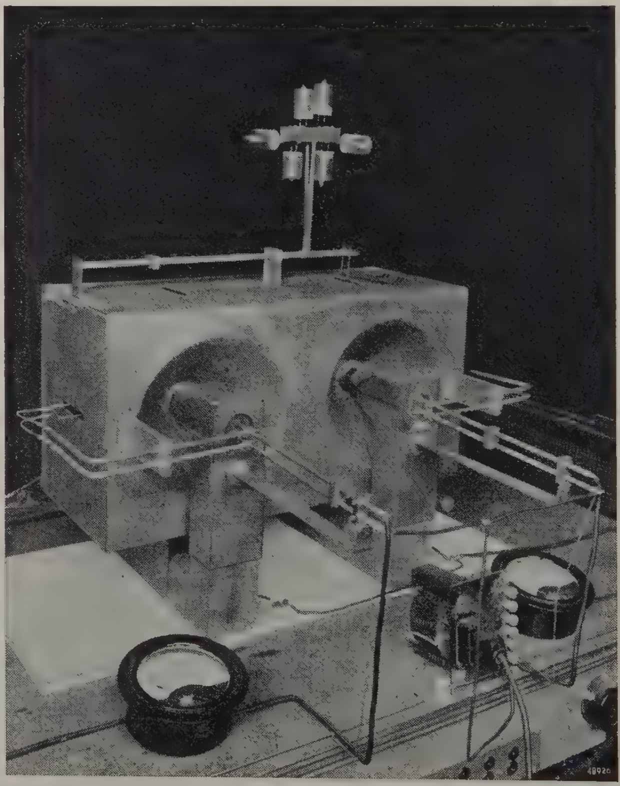

Fig. 10. An induction-tube oscillator of high power. The two induction tubes are mounted in the cavity resonator at the point of voltage maxima. In this way use is made of the fact that at higher forms of oscillation a cavity resonator has more than one voltage maximum.

If only one characteristic oscillation is excited the quality factor Q can be defined by the following relation:

$$Q = 2\pi \frac{U}{W_T}, \quad . \; . \; . \; . \; . \; . \; . \quad (13)$$

in which U is the average field energy and W_T the heat developed during one period. Definition (13) is very generally valid, not only for ordinary oscillation circuits (*i.e.* with concentrated L and C) but also for cavity resonators. If a number of characteristic oscillations are excited simultaneously the quality factor for each characteristic oscillation is also given by (13), provided two characteristic oscillations are never so close to each other that the peaks of the resonance curve coalesce. Equation (13) does not, for example, hold in the case of a cylinder with an inner partition (see fig. 14) to be discussed later.

The significance of definition (13) becomes clearer when it is borne in mind that the average field energy of a free oscillation with frequency ω decreases owing to losses in the course of the time t according to the formulae

$$U(t) = U\, e^{-at},$$

where a is the damping constant. If a is small the energy lost during one period is given by

$$W_T = U(0) - U\left(\frac{2\pi}{\omega}\right) = U\left[1 - e^{-2\pi a/\omega}\right] \approx a\frac{2\pi}{\omega}\,U.$$

By comparison of this result with (13) we find that

$$Q = \frac{\omega}{a}.$$

The quality factor thus actually becomes larger according as the resonance of the cavity resonator is sharper, because the sharpness of resonance of an arbitrarily oscillating system increases when the damping constant decreases.

In order to be able to calculate the quality factor, one must, according to (13), know the average field energy and the energy dissipated during one period.

If A is the magnetic field strength, the field energy is proportional to the integral of H^2 over the whole volume of the cavity resonator. The energy dissipated is given by the Joule heat of the currents induced by H in the walls of the cavity resonator (we assume here that the radiation losses may be disregarded). Now the density of the induced current at every point of the wall is proportional to the magnetic field strength at that point and the resistance of the wall is proportional to the depth of penetration δ due to the skin effect [6]). The energy dissipated is therefore finally proportional to δ and to the integral of H^2 over the whole surface of the cavity resonator. The calculation gives indeed

$$Q = \frac{2}{\delta}\,\frac{\iiint H^2\,d\tau}{\iint H^2\,d\sigma}, \quad . \; . \; . \; . \quad (14)$$

when $d\tau$ and $d\sigma$ are volume and surface elements of the cavity resonator respectively. In many cases the order of magnitude of Q is given in sufficient approximation by

$$Q = \frac{2}{\delta}\,\frac{\text{volume of the cavity resonator}}{\text{surface of the cavity resonator}}. \quad (15)$$

From (15) it follows immediately that for "three-dimensional" cavity resonators Q will in general be larger than for flat cavity resonators. For the latter we have in mind especially the square plane cavity resonators — it may be concluded from (15) that for higher forms of oscillation Q has about the same value as for the fundamental oscillation, if λ, and consequently δ, which is proportional to $\sqrt{\lambda}$, are kept constant, *i.e.* when cavity resonators of different, suitably chosen dimensions are compared with each other. Such higher forms of oscillation were indeed obtained by placing the cavity resonators side by side and then removing the partitions between them; and for really thin cavity resonators the heat development in the side walls amounts to almost nothing compared with that of bottom and cover.

For the cube Q is equal to $2\,a/3\delta$, where $2a$ is the length of an edge. Since the wavelength λ of the fundamental oscillation is equal to $2a\sqrt{2}$ and the depth of penetration for copper is equal to $4 \cdot 10^{-5}\sqrt{\lambda}$ (λ in cm), one finally obtains for the quality factor Q of a copper cube oscillating on the longest wavelength $Q = 99000\sqrt{a}$ (a in cm).

We must now consider briefly the resonance resistance of a cavity resonator. We define here the resonance resistance for flat cavity resonators only, because it is then still possible to speak of a voltage V, as explained above. The resonance resistance is a quantity Z such that the heat development W in the cavity resonator per unit of time is given by $W = V^2/Z$.

Since the voltage for flat cavity resonators is a function of position, the resonance resistance will in general also depend upon position. Usually by resonance resistance without any indication of position is meant the resonance resistance at the voltage maximum.

---

[6]) If ϱ is the specific resistance of the material of the wall of the cavity resonator, the resistance of the layer in which an appreciable current flows is proportional to ϱ/δ. It now follows from the theory of the skin effect that δ is proportional to $\sqrt{\varrho}$. The resistance of the layer in question is therefore indeed proportional to δ.

We shall now examine how the resonance resistance for higher forms of oscillation behaves compared with that for the fundamental oscillation in the case of square plane cavity resonators, again at constant λ. When n similar square cavity resonators, each of which executes the fundamental oscillation, are placed side by side, it is clear that n times as much heat is developed as when only one of them is oscillating. If we now remove the intermediate partitions, as described above, and consider the whole as a single cavity resonator in a higher state of oscillation, its resonance resistance is thus n times as small as that of the elementary cavity resonator of which it was built up (we again disregard the heat development in the partitions).

The advantage of the use of the concept "resonance resistance" for cavity resonators is often manifested in the fact that in many arrangements the voltage which will act on a cavity resonator can be calculated as soon as the resonance resistance is known.

We shall not, however, go into that here.

Coupled cavity resonators

We shall now discuss the lowest characteristic oscillations of a cylindrical cavity resonator divided into two by a transverse partition midway along its length; the partition is provided with a circular hole exactly in the centre.

Such a cavity resonator may be considered in two ways: on the one hand as the limiting case of a cylinder with constriction in the middle (*fig. 11a*), and on the other hand as two separate resonators coupled by a small hole in the partition forming the "cover" of one and the "bottom" of the other (fig. 11*b*). As long as the partition is entirely closed there may be quite independent oscillations, with any arbitrary phase difference, on either side of the partition. When there is a hole that is no longer true, for instead of the characteristic frequency of the fundamental oscillation there are now two characteristic frequencies, at one of which the oscillations to the left and right are in the same phase and at the other in opposite phases.

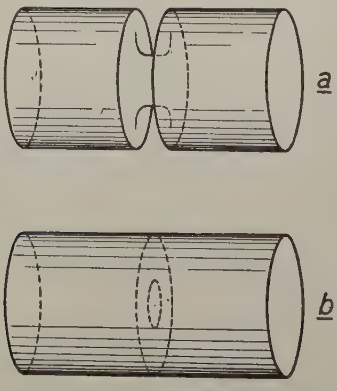

Fig. 11. A cylindrical cavity resonator divided into two by a transverse partition midway along its length. There is a small round hole at the centre of the partition. Such a cavity resonator may be considered: *a*) as a single cylindrical cavity resonator with a constriction at the middle; *b*) as two separate cavity resonators coupled by the hole.

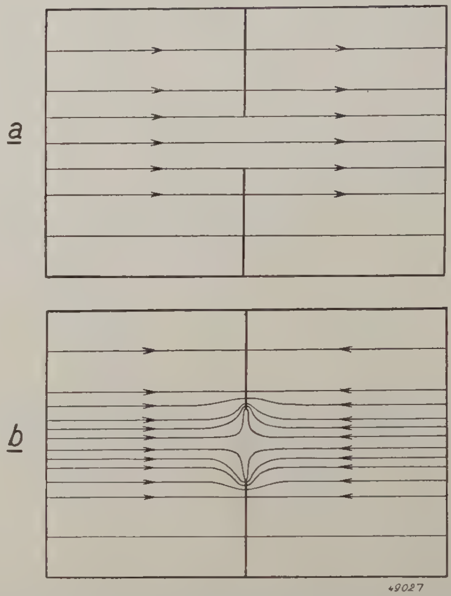

Fig. 12. The course (diagrammatic) of the electric lines of force for the two lowest characteristic frequencies of the cavity resonator of fig. 11. (The forms of oscillation here having axial symmetry, it is thus sufficient to show a single cross-section.) These characteristic frequencies are both formed from the characteristic (fundamental) oscillation which would be present if there were no hole in the partition.

In the first case, that of the same phase, thus with the electric field similarly directed on the left and right — for the fundamental oscillation the electric field in a cylinder closed at both ends is everywhere parallel to the axis — the situation is always the same with or without partition, with or without a hole in it. Thus when a hole is made in the partition there is no difference in the case of this characteristic oscillation, the characteristic frequency ω_0 of the cavity resonator remaining unchanged and also the course of the lines of force (*fig. 12a*).

It is different in the second case where the oscillations to the left and right of the partition are in opposite phase. Lines of force can indeed end on a conductor, but not somewhere in free space. Thus difficulties are encountered at the hole and we must find out what happens then. Actually the lines of force will curve in the vicinity of the hole, so that they end on the partition or at least near the edge of the hole (fig. 12*b*). The course of the lines of force is thus slightly different from that in the case with no hole in the partition. This is not without consequences for the value of the characteristic frequency. It would take us too far afield to explain how, after having obtained a picture of the variation of the field by means of the above reasoning, by means of a relatively simple calculation a formula can be found giving the frequency correction, at least for a very small hole. The formula in question is as follows:

$$\omega'' = \omega_0 \left\{ 1 + 0.788 \left(\frac{\varrho}{R}\right)^3 \frac{R}{l} \right\}; \quad . \quad . \qquad (16)$$

where ω'' is the frequency caused by the presence of the hole, ω_0 the "undisturbed" frequency, R and $2l$ are respectively the radius and the total length of the cylinder and p the radius of the hole.

Thus when the hole is bored in the partition the characteristic frequency ω_0 of the cavity resonator is split into two: $\omega' = \omega_0$, and ω (given by equation (16)). For a very small hole ω' and ω'' lie very close together, so that the peaks of the resonance curve corresponding to ω' and ω'' partially coalesce, exactly as in the case of two coupled circuits with concentrated L and C. And just as in the case of two such coupled circuits, the familiar resonance curve with two "shoulders" will occur here as soon as there is any damping in the cavity resonator. In *fig. 13* a photograph is shown of such a resonance curve for a cavity resonator with partition; this

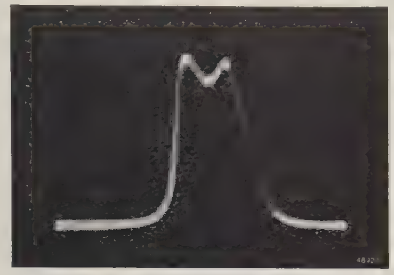

Fig. 13. Resonance curve of coupled cavity resonators recorded with the apparatus of fig. 14. The frequency is here "plotted" in the horizontal direction. The deviation in the vertical direction is a measure of the intensity of the oscillation excited in the cavity resonators.

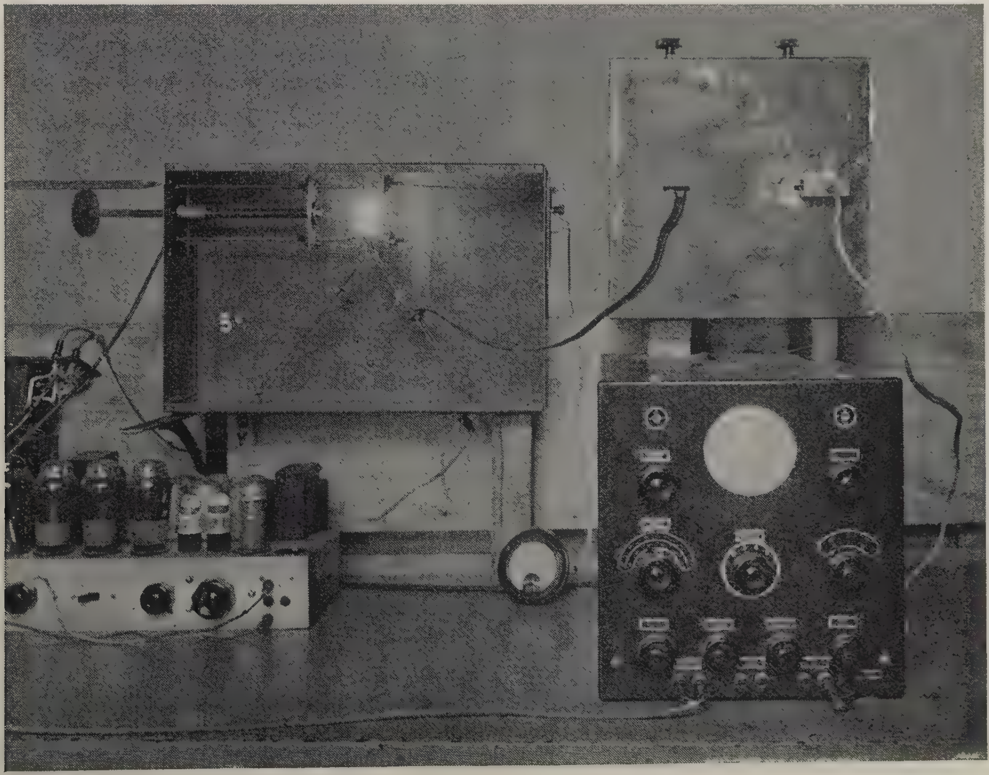

Fig. 14. Apparatus for the recording of the resonance curve of cavity resonators. The resonator visible here on the upper right is divided into two by a partition with a hole in it (the partition is vertical and perpendicular to the plane of the figure), *i.e.* in this case it is actually a question of two coupled cavity resonators. The oscillations are excited with the help of the loop, part of which may be seen protruding from the slot on the left of the cavity resonator, which loop is connected with the oscillator on the left in the figure, whose frequency can be varied. The high-frequency alternating current excited in the loop situated in the other slot is applied to the vertical plates of the oscillograph after rectification. A separate arrangement provides that the horizontal deflection of the electron beam in the oscillograph will be proportional to the variation of frequency.

resonance curve was recorded with the apparatus shown in *fig. 14*. (In the case in question the cavity resonator was square, which of course does not alter the qualitative aspect of the phenomena.) It is thus understandable that the "composite" cavity resonator considered here has an analogous function in high-frequency technology to that of coupled circuits in the region of lower frequencies, in particular to that of a band filter.

Equation (16) was valid for a very small hole ($\varrho \ll R$). What will the frequency correction be when the hole is no longer very small? It is very difficult to answer this question satisfactorily. One thing is certain, however: when the hole is so large that there is no partition left at all, *i.e.* when $\varrho = R$, the two characteristic frequencies ω' and ω'' must clearly be characteristic frequencies of an undivided cylinder of radius R and length $2l$. The theory of the characteristic frequencies of such a cylindrical cavity resonator shows that the course of the lines of force for the lowest (rotation-symmetrical) characteristic oscillations is as represented in *figs. 15a*, *b* and *c*. We compare these figures with fig. 12*a*, *b* drawn for the case of a small hole. Since the course of the lines of force varies continuously with the size of the hole, fig. 12*a* must correspond to fig. 15*a* and fig. 12*b* to fig. 15*b*. It should therefore be possible to obtain the frequency corresponding to fig. 15*b* from the frequency corresponding to fig. 12*b* when ϱ is increased continuously. The value of the characteristic frequency for the forms of oscillation in fig. 15*b* is familiar from the theory:

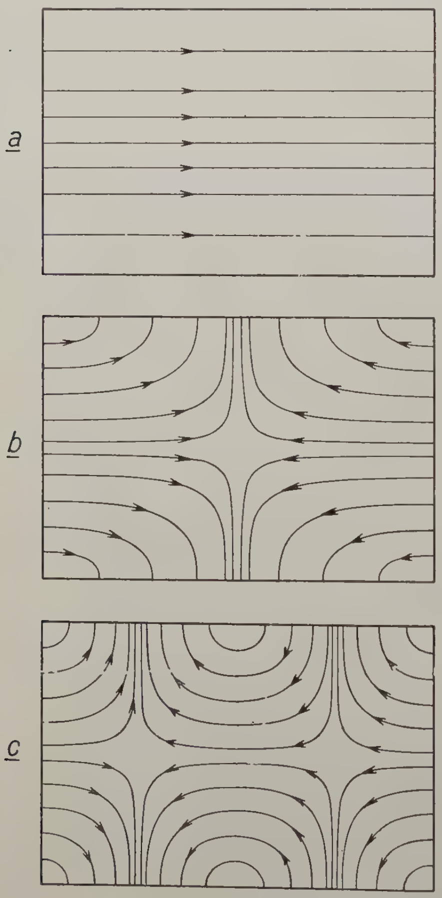

Fig. 15. The (diagrammatic) course of the electric lines of force for the lowest characteristic frequencies of a cylindrical cavity resonator; the forms of oscillation in question are axial symmetrical: *a*) corresponds to the fundamental oscillation; *b*) and *c*) to the two succeeding characteristic oscillations.

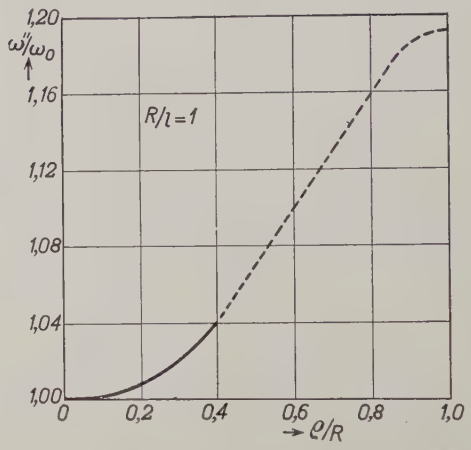

Fig. 16. The lowest but one characteristic frequency ω'' plotted vertically of the cavity resonators coupled by a hole as a function of the ratio ϱ/R (plotted horizontally), of the radius ϱ of the hole and the radius R of the cylinder. The dotted part of the curve is not well known theoretically.

$$\omega''_{\varrho=R} = \omega_0 \sqrt{1 + 0.426\frac{R^2}{l^2}} \quad . \quad . \quad . \quad . \quad . \quad (17)$$

Of the curve which gives ω'' as a function of ϱ/R we now have a section at the beginning, given by equation (16), and at the end point ($\varrho/R = 1$), given by equation (17) — see *fig. 16*. We shall now guess at the rest by drawing the dotted part of the curve in a reasonable manner. The frequency corresponding to figs. 12*a* and 15*a* is independent of ϱ and l, namely in both cases equal to ω_0, *i.e.* it will be represented in fig. 16 by the ϱ/R axis [7]).

We may not conclude without drawing attention to the analogy which exists between the electromagnetic oscillations in Lecher systems, flat and general cavity resonators on the one hand and on the other the acoustic vibrations of strings, membranes and Helmholtz resonators. This analogy is no identity; there are certain differences between the two groups of phenomena. But the expectation, which is perhaps natural, that there must exist an electromagnetic counterpart of the speaking-tube and the horn is actually confirmed by the electromagnetic "wave guides" and "horn aerials", which we shall not, however, discuss in this article.

---

[7]) *Cf.* H. A. Bethe, Phys. Rev. **66**, 163, 1944.

THE PRACTICAL CONSTRUCTION OF VIBRATION-FREE MOUNTING WITH AUXILIARY MASS

by J. A. HARINGX.

621-752

In order to reduce the amplitudes of the forced vibrations of an instrument due to the motion of its surroundings it is usually mounted on a resilient construction which at the same time is provided with a certain damping in connection with the decay of the free vibrations. In a previous article the advantages were shown of introducing the damping between the apparatus and an auxiliary mass attached to it with springs. The features of this system were discussed for the one-dimensional case. In the present article the case of multi-dimensional vibrations is examined (translations parallel to and rotations about various axes). For such a case the mounting should preferably satisfy certain conditions of symmetry. For the resilient support practically only helical steel springs can be considered, use being made not only of their axial rigidity but also of their rigidity with respect to lateral deflection. Further consideration is also given to the practical manner of introducing the damping.

The use of sensitive instruments such as balances, galvanometers and microscopes is often made difficult or even impossible by vibrations of the surroundings. In order to reduce the amplitudes of the forced vibrations of such an instrument due to the vibrations of the surroundings it is placed upon sufficiently weak springs. It is then necessary, however, to introduce a damping to ensure that the decay of the free vibrations of the system caused by slight impulses or initial displacements is rapid enough. In a previous article [1]) we showed the advantages of introducing this damping between the apparatus and an auxiliary mass attached to it with springs. The features of this system were discussed, but only for the case of a one-dimensional movement.

When it comes to putting these ideas into practice we are immediately faced with the problem that the foundations of an apparatus are, in general, apt to execute vibrations in and around various directions. Since in most cases an instrument is sensitive to several of these vibrations, the system of mounting must provide for resiliency in several directions. This cannot be done, however, in an arbitrary manner as regards the choice of the centres of the elastic forces and their directions, for then it is quite likely that once the apparatus begins to vibrate it will show highly complicated and absolutely indeterminable oscillations. It would then be impossible to determine what measures have to be taken to restrict the forced vibration and to stop as quickly as possible the free vibrations of the system after an impulse or an initial displacement. We shall now explain very briefly how this problem has to be dealt with.

A rigid body suspended in space by a set of springs possesses six degrees of freedom, *viz:* translation in three mutually perpendicular directions and rotation about three mutually perpendicular axes. These degrees of freedom are as a rule coupled, that is to say a force acting in one of the directions (for instance the reaction of the springs or the inertia of the mass) results not only in a translation in this direction but also in other translations and rotations. This is the reason why the behaviour of the vibrating system is so complicated. Under certain conditions however the coupling between the different degrees of freedom may disappear, namely when the spring mounting possesses a so-called centre of elasticity with three perpendicular principal main axes of elasticity passing through this centre and coinciding with the principal axes of inertia of the suspended body (the centre of gravity then automatically coincides with the centre of elasticity). A principal axis of elasticity is defined in such a way that a force acting in its direction causes a translation in this direction only, whilst a couple about such an axis likewise causes a rotation only about that axis, so that under the conditions mentioned there are in fact six mutually non-coupled degrees of freedom. Any arbitrary combination of elastic attachments will not as a rule possess a centre of elasticity. Such will be the case, however, if for instance the mounting is symmetrical with respect to two mutually perpendicular planes, like those in *figs. 1a-c*, and if at the same time the elastic elements are of equal rigidity in the directions perpendicular to these planes. For reasons of symmetry the intersecting line of these two planes, *i.e.* the z axis, must be one of the principal axes of elasticity and the two other principal axes must be parallel to the x and y axes. The height of the centre of elasticity on the z axis can be calculated from the rigidity of the springs, an example of which will be given below.

---

[1]) J. A. Haringx, Vibration-free mountings with auxiliary mass, Philips Techn. Rev. 9, 16, 1947.

If, now, in our vibration-free mounting the springs are fitted in this two-fold symmetrical manner and care is taken that the principal axes of inertia of the mounting combined with the instrument placed on it coincide with the principal axes of elasticity, *i.e.* the x, y and z axes, then the vibration of the system following the six degrees of freedom are approximately [2]) independent of each other and can be treated as six separate one-dimensional problems. This also holds when applying the principle of our vibration-free mounting sketched in *fig. 4* of the previous article [1]) by attaching an auxiliary mass to our apparatus by means of a number of elastic and damping elements. This attachment must also have three principal axes of elasticity coinciding with the same x, y and z axes, the principal axes of inertia of the auxiliary mass likewise coinciding with these axes. Moreover, the damping elements must be introduced in such a way that in case of translations parallel to or rotations about the three said axes the damping results in forces along and couples about the same axes.

When these conditions are satisfied — in practice this can be done with sufficient accuracy — we can apply the results of our previous study of the choice of the parameters to each degree of freedom separately and thus endeavour to keep the system free of all disturbing vibrations.

The design of the springs

The mechanical disturbances apt to give trouble in a building may be caused by the regular throbbing of a machine, but they may also be due to all kinds of irregular shocks such as footsteps, the slamming of doors, etc. Consequently any frequencies, even the very lowest, may be represented in the Fourier spectrum of the disturbances. The primary effect of the disturbances is to set in motion the floors, walls and tables in the building, which then execute their natural vibrations. The "vibration systems" of floors, tables, etc. act as a kind of filter, which preferentially transmits the vibrations in the neighbourhood of the resonant frequencies. We shall thus need to analyse the disturbing vibrations more closely with the aid of a vibration pick-up. It has been found that the lowest of the natural frequencies mentioned usually lie between 15 and 20 c/s. If, by means of a vibration-free mounting, the amplitudes of an apparatus are to be reduced, for example, to a few percent of those of the foundation, then under these conditions we must give the mounting the very low resonant frequency of 2 to 4 c/s [3]).

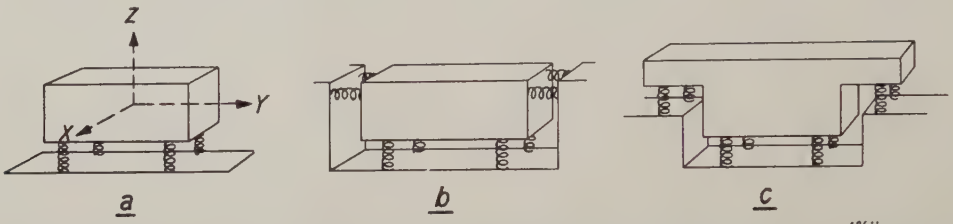

Fig. 1. Three different arrangements for making an apparatus vibration-free where multi-dimensional disturbing vibrations exist. The mountings are all symmetrical with respect to two perpendicular vertical planes.

Such low resonant frequencies preclude the use of rubber cylinders for the resilient elements of the mounting, because they are too rigid. For the resonant frequency ω_0 of a vibrating system we may write

$$\omega_0 = \sqrt{\frac{g}{f}},$$

where g is the acceleration due to gravity and f the compression of the resilient element caused by the weight of the mass. In order to obtain, for example, a resonant frequency of $\omega_0/2\pi = 3$ c/s, f must therefore be 2·8 cm. If rubber cylinders were to be used as spring elements their length l would have to be at least 10 cm to permit such a large compression. Moreover, to prevent them from buckling, the diameter D would have to be of the same order of magnitude, so that the force necessary for the compression:

$$P = \pi D^2 \cdot E \cdot f/l$$

in the case of the weakest variety of rubber with a modulus of elasticity $E = 10$ kg/cm², would be of the order of 1000 kg. A mounting with four such cylinders should thus have a weight of 4 tons! Obviously such heavy constructions cannot generally be used as a support for a balance or a microscope. A reasonable total weight would be say 50 to 100 kg. Such a weight could be more closely approached if sponge rubber were used, but as this material is not very durable it is not to be recommended for our purpose.

---

[2]) "Approximately" because the force of gravity exercises a disturbing influence.

[3]) This result is obtained by requiring, in formula (8) of the article referred to in footnote [1]), that for instance $\left|\frac{a}{a_0}\right| =$ 0·04 and setting μ equal to 0·5, so that $\omega/\omega_0 = 7$ (ω = given disturbing frequency, ω_0 = resonant frequency to be chosen).

As resilient elements for a vibration-free mounting helical steel springs are much more suitable. These can easily be made so as to take up the desired compression of about 3 cm under a force of about 10 kg. Helical steel springs possess a large number of distinct natural frequencies, with the result that the vibrations having corresponding frequencies are transmitted to the instrument. This fact is in contradiction to the statements previously given and is often considered as a serious objection, but fortunately these natural frequencies usually lie so high that — at least in our case — they do not cause any trouble at all.

As a rule it is only the axial rigidity of such springs that is utilised; take for instance buffers, spring balances, etc. But, in addition to this axial rigidity, springs also possess a lateral rigidity, provided their ends are not both hinged. Therefore, since we have to do with a three-dimensional mounting having six degrees of freedom it is obvious that the construction can be simplified by turning this lateral rigidity to account (see figs. 1*a* and 1*c*).

Let us consider for example the arrangement sketched in *fig. 2*, which may serve as a prototype for many practical applications. A flat plate *p*, on which is placed the balance or other instrument to be rendered free of vibration, is supported by four helical springs. The auxiliary mass, a second plate *P*, is supported by the first plate in a similar manner. In both cases the four springs simultaneously perform the function of a resilient attachment in the vertical direction and in the two horizontal directions.

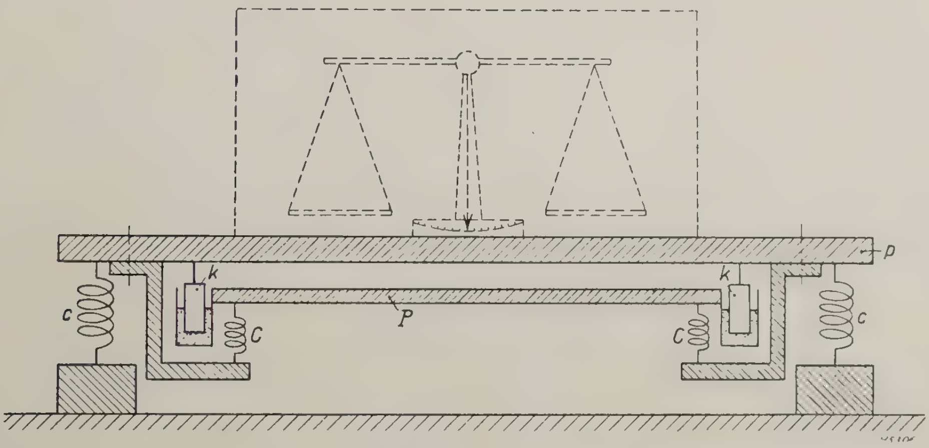

Fig. 2. Model of the practical construction of a vibration-free mounting with auxiliary mass. The plate *p* supported by the helical springs *c* and carrying the instrument represents the main mass, the plate *P* connected below with the helical springs *C* is the auxiliary mass. Between these two plates the damping *k* has been provided.

The question arises as to how the springs must be designed in order to get a certain lateral rigidity in addition to a given axial rigidity. This axial rigidity, *i.e.* the force required per unit of axial compression, can be calculated with the help of the formula

$$c = \frac{4m}{m+1} \, \frac{EI}{\pi n D^3}.$$

(E = modulus of elasticity, I = moment of inertia of the circular cross section of the wire, n = number of turns, D = coil diameter, m = lateral contraction or Poisson's ratio; for steel $m = 10/3$). For the calculation of the lateral rigidity, however, the theory of the lateral deflection of helical springs had first to be worked out in more detail [4]), and at the same time the danger of buckling could also be studied. Here we can give the most important results only very briefly.

A slender helical spring will buckle as soon as the compression load, and thus also the relative compression ξ, has reached a certain value. The critical compression ξ_k is found to depend — at least in a first approximation [5]) — only on the ratio of the length l_0 of the unloaded spring and its diameter D, the so-called ratio of slenderness of the spring. The relation found for the case where both ends of the spring remain parallel after deflection (as in fig. 2) is shown in *fig. 3*. It is seen that when $l_0 < 2{\cdot}6\ D$ the spring can never buckle. This condition must therefore be taken into account when designing the springs.

---

[4]) J. A. Haringx. On the buckling and the lateral rigidity of helical compression springs. Proc. Kon. Ac. v. Wet Amsterdam *45*, 533-539 and 650-654, 1942.

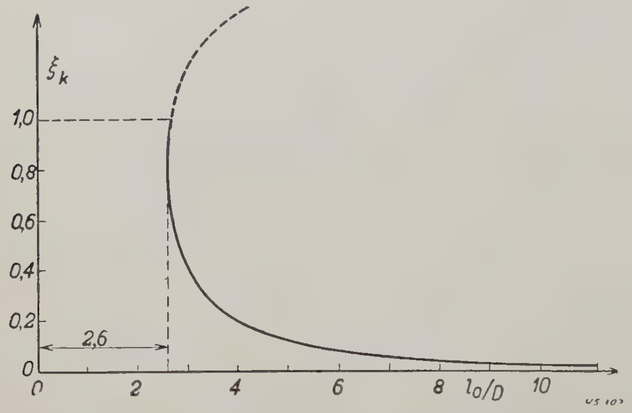

Fig. 3. A loaded helical spring without axial guidance will buckle under a certain force corresponding to a certain (critical) compression ξ_k. ξ_k is found to depend only on the ratio of slenderness l_0/D of the spring. From the relation shown here, which holds for a spring with its ends both hinged or remaining parallel, it may be seen that with a ratio of slenderness $l_0/D < 2{\cdot}5$ buckling can never occur. When both ends of the spring are fixed l_0 is to be taken as half the length of the spring.

Further, the lateral deflection y of the free end of an axially compressed spring under the simultaneous influence of a lateral force L and a couple M (*fig. 4*) is given by the equations

$$\left.\begin{aligned} L &= c_1 y - c_2 \psi, \\ M &= -c_2 y + c_3 \psi, \end{aligned}\right\} \quad . \; . \; . \; . \; . \quad (1)$$

where ψ is the rotation of the upper end and c_1, c_2 and c_3 are the coefficients of rigidity. These coefficients are related to the axial rigidity c according to the following equations:

$$\frac{c}{c_1} = \frac{m}{2\,(m+1)}\left[1 + \frac{2m+1}{3m}\left(\frac{l}{D}\right)^2\right]\cdot \varepsilon_1, \quad . \; (2a)$$

$$\frac{c_2}{c_1} = \frac{1}{2}\; l \cdot \varepsilon_2, \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; (2b)$$

$$\frac{c_3}{c_1} = \frac{mD^2}{4\,(2m+1)}\left[1 + \frac{4\,(2m+1)}{3m}\left(\frac{l}{D}\right)^2\right]\cdot \varepsilon_3. \; . \; (2c)$$

Here l is the length of the spring when loaded and ε_1, ε_2, ε_3 are correction factors depending upon the relative axial compression ξ and the ratio of slenderness l_0/D. These three factors are represented graphically in *figs. 5a-c*.

By applying the results described to the mounting in *fig. 2* we can first of all calculate the height h of the centre of elasticity above the base plane of the springs.

A horizontal force L through this centre — according to the definitions already given — will cause only a lateral displacement y of the plate; the plate remains horizontal, thus $\psi = 0$. It follows from (1) that in this case the couple M caused by L must be:

$$M = -\frac{c_2}{c_1}\cdot L,$$

while on the other hand this couple is given by

$$M = -(l - h)\cdot L.$$

By eliminating M/L from these two equations and making use of (2b) we obtain:

$$h = l\,(1 - {}^1/_2\,\varepsilon_2).$$

The centre of gravity of the movable mass of the mounting must therefore lie at this height. Further, under this condition, we have $L = c_1 y$, so that equation (2a) gives us directly the rigidity of the spring c_1 required to make our mounting vibration-free as regards transverse vibrations.

It will often be desired to have the mounting behave in the same way in the vertical as in the horizontal direction. When vertical springs alone are used c_1 must then be equal to c. This establishes a relation between ξ and l_0/D which can be derived

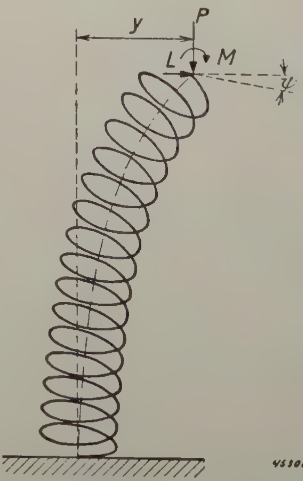

Fig. 4. When a transversally directed force L and a couple M act on the free end of a helical spring whose other end is clamped and which is compressed axially by a weight P, a lateral deflection y and a rotation ψ occur at that free end.

---

[5]) A more precise calculation shows that the pitch and the wire diameter exercise a certain, although slight, influence. This calculation will be published elsewhere in a paper in which other problems discussed in this article are also worked out in more detail.

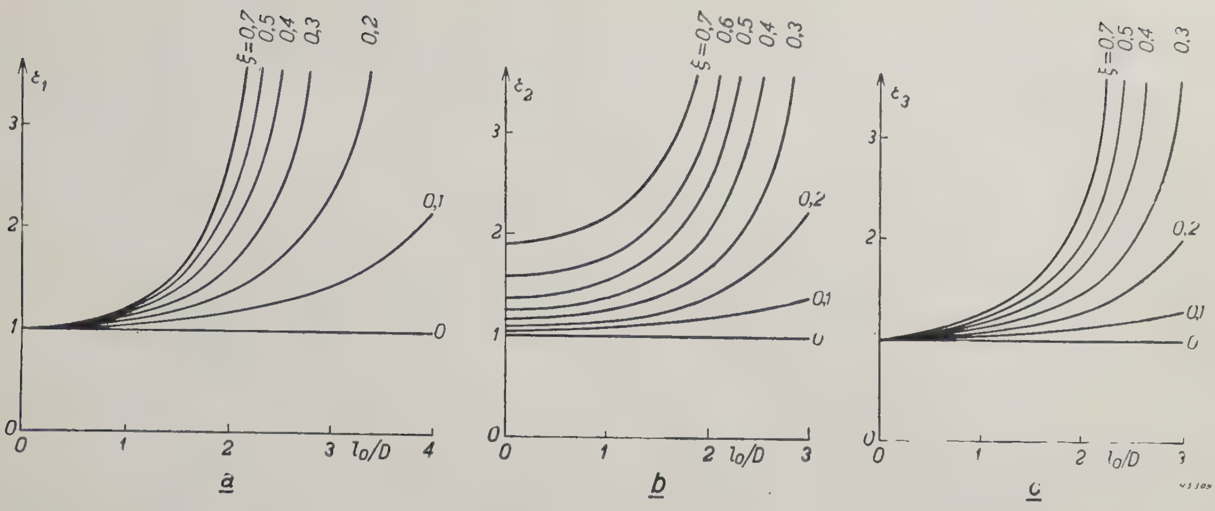

Fig. 5. The correction factors ε_1, ε_2, ε_3 occurring in formulae (2) as functions of the ratio of slenderness l_0/D with the compression ξ as parameter.

from equation (2a) and fig. 5a. This relation is represented graphically in *fig. 6*. With the aid of this graph and the prescribed loaded and characteristic frequencies it is possible to derive all the conditions that have to be satisfied by the parameters of the spring (length, diameter, number of turns, etc.).

Also the rigidity of the mounting with respect to a couple about one of the principal axes of elasticity can be expressed in the quantities c, c_1, c_2, c_3; in this case the contributions of the rigidities c_1, c_2 and c_3 can often be disregarded. For counteracting the rotation vibrations connected with this couple the same considerations are valid as were given for translation vibrations, except that the various moments of inertia take the place of the masses. There is nothing new of any importance in this, so that we shall not go deeper into it here.

The practical application of the damping

The damping can best be realised by means of four cups k filled with oil of a certain viscosity and affixed to the auxiliary mass (plate P in fig. 2), in which cups four others affixed to the main mass (plate p) can move freely in all directions with a clearance of some millimeters.

When it is desired that the damping coefficient k should be the same in the horizontal and in the vertical directions the damping cups must be square and, as calculation has shown, they must be filled to a height $H \approx 1.45\ B$, where B is the side of the basal plane of the inner cup. The damping coefficient in that case is approximately equal to

$$k \approx 18\,\frac{\eta B^4}{d^3}\left(1 + 2.5\,\frac{d}{B}\right),$$

where d is the distance between the walls of the two cups and η the viscosity of the liquid.

Eventually the optimum damping can be adjusted experimentally by using oils of a different viscosity, but there is really no very great sensitivity in such an adjustment, as we have seen in the article already referred to [1]). Even if the adjustment is not correct the free vibrations after an impulse die out quickly enough and the forced vibrations of the system due to the motion of its surroundings are scarcely affected at all. It is really fortunate that this adjustment is not critical, for the viscosity of the oil — and thus its damping power — depends closely upon the temperature, so that we can obtain the optimum behaviour only at one temperature.

Instead of the liquid damping that we have so far assumed to be applied, a frictional damping can also serve to dissipate the kinetic energy imparted to the mounting by an impulse; for instance a set of flexible strips can be attached to one of the plates in fig. 2 with a slight pressure

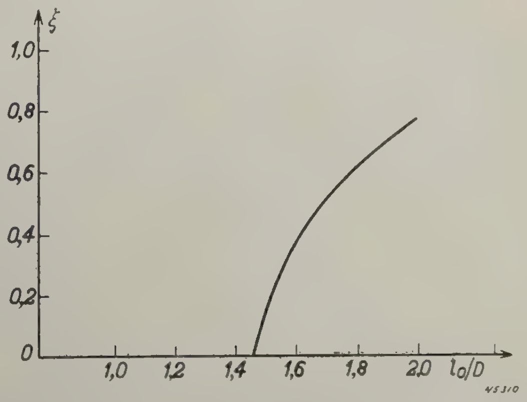

Fig. 6. The relation between the compression ξ and the ratio of slenderness l_0/D which follows from the condition that the spring must have the same axial and lateral rigidity.

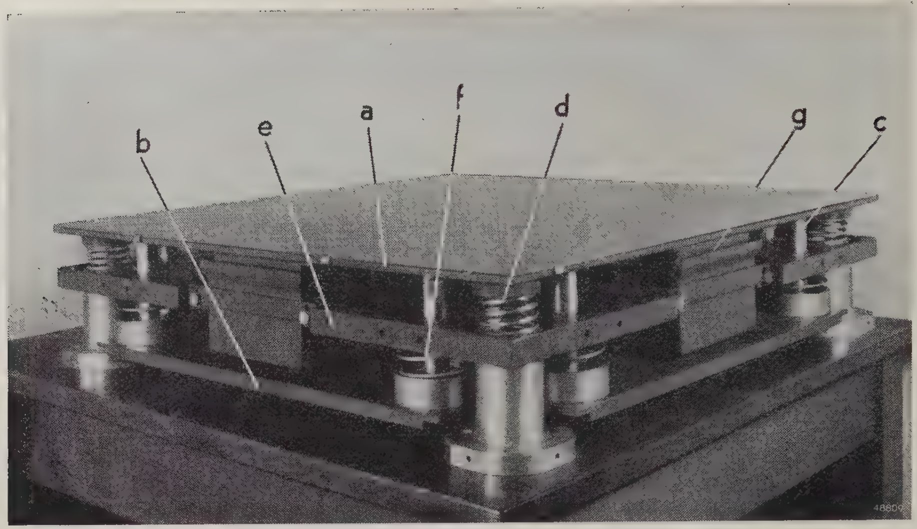

Fig. 7. Practical construction of a vibration-free mounting. The dimensions of this apparatus are 60 × 60 × 15 cm; its weight is 110 kg. The two plates *a* and *b* together form the main mass. They are connected with twelve rods *c*. This main mass rests upon four helical springs *d*. The auxiliary mass *e* is coupled to the main mass with eight springs *f*. The cup *g* is one of the four damping elements.

against the other plate and rubbing over it as soon as a relative displacement takes place. When the amplitudes are sufficiently large it is possible with this very simple construction to obtain a fairly rapid dissipation of the kinetic energy. When, however, the amplitudes are less than a certain small value a more or less unexpected phenomenon occurs; the frictional forces cause one mass to follow the movements of the other. Thus, as soon as the frictional damping has caused the amplitude to drop to this small value it ceases to act and for some time the instrument will continue to vibrate at this amplitude. If, as is usual, the resonant frequency is low, this may not eventually constitute any objection for some instruments, such as dial gauges, microscopes and not too sensitive balances, but for galvanometers and other highly sensitive instruments it will be necessary to use liquid damping.

In *fig. 7* a vibration-free mounting is shown which is in use in this laboratory and which was designed according to the theory developed here. This apparatus, weighing 110 kg, can be used as a mounting for instruments up to 35 kg. The resonant frequency (with rigidly coupled main and auxiliary masses) lies at about 3 c/sec. This arrangement fully answers expectations both as regards the decay of the free vibrations and as regards the amplitudes of the forced vibrations due to the motion of the surroundings.

A REMARKABLE PROPERTY OF TECHNICAL SOLID DIELECTRICS

by M. GEVERS and F. K. du PRÉ.

537.226.3:537.226.8

In this article a remarkable property of technical solid dielectrics is discussed. The ratio between the temperature coefficient of the dielectric constant ε and the quantity tan δ (δ = loss angle) is practically constant for most materials. Expressed in a formula this is $(1/\varepsilon)\,\dfrac{\partial\varepsilon}{\partial T} = A \tan\delta$, where $0{\cdot}04 < A < 0{\cdot}09$ and usually $A \approx 0{\cdot}05$. An explanation of this and other properties is given, viewed from the same aspect. A few exceptions to the above rule are described and explained, while some practical conclusions are drawn.

In electrotechnics solid dielectrics are used chiefly in two ways, namely:

1) for the insulated attachment and leading in of supply leads, and
2) as media in condensers.

For these purposes many different materials are used which may be divided into the following groups according to their structure and chemical composition:

a) natural dielectrics, such as amber, mica, quartz;
b) inorganic materials such as glass, quartz glass, porcelain, steatite (Mg-silicate), cordierite (MgAl silicates), rutile (TiO_2) and related products, Mg titanates;
c) organic materials such as condensation products (phenol-formaldehyde resin), polymerization products (polystyrene), rubber and ebonite.

The properties which are required of these dielectrics depend upon the particular use to which they are put. In every case more or less good insulation is required. In case 2) a high value of the dielectric constant ε is usually also desired.

Another requirement usually made when such materials are used in connection with alternating current technics is that the dielectric should have small dielectric losses. In case 1) the reasonableness of this requirement is obvious, since losses lead to undesired temperature increase and finally to breakdown and destruction of the material. But also in case 2) low dielectric losses are usually desired, namely when the circuit in which the dielectric is used between the electrodes of a condenser has to have a high selectivity. In a previous article [1], in which dielectric after-effect phenomena were discussed, it was already pointed out that losses have an unfavourable effect on selectivity. This becomes clear when it is kept in mind that the influence of these losses corresponds to the inclusion of a resistance in the *LC*-circuit.

In many cases a second requirement must also be made of an oscillation circuit, namely that the frequency of the circuit shall be more or less independent of the temperature of the surroundings. In this case, therefore, it will be required of the medium that the temperature coefficient of the dielectric constant shall be small.

From the above it is evident that for various technical purposes it is of importance to measure not only the losses (determined by the so-called loss angle δ) but also the temperature coefficient of the dielectric constant (for the sake of brevity called from now on simply "temperature coefficient"). Since in many respects the data from literature were incomplete, special attention has been devoted in the Philips laboratory to the determination of these two properties with the greatest possible accuracy [2]. A remarkable relation between the two quantities has thereby been brought to light, namely that, apart from a few exceptions to be mentioned later, the substances which have a high temperature coefficient have at the same time a large loss angle, while the ratio between the two quantities is practically constant.

This can be understood theoretically when bearing in mind that the losses as well as the temperature coefficient are affected by so-called after-effect phenomena.

In this article some numerical data on the properties in question will be given. Further, beginning with certain conceptions about the cause of the after-effect (namely dipoles which are orientated with a time lag in a field) expressions will be derived for the loss angle and the temperature coefficient, in which the proportionality mentioned appears, while the value of the proportionality factor proves to correspond to the value found experimentally.

Dielectric losses

Let us suppose that a dielectric is situated in a constant electric field *E*. Under the influence of that field a dielectric displacement D occurs where

---

[1] J. L. Snoek and F. K. du Pré, Philips Techn. Rev. 8, 57, 1946.

[2] *Cf.* M. Gevers, Philips Res. Reports, 1, 197, 279, 361, 447, 1946.

$$D = \varepsilon E \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

ε being the dielectric constant.

If the dielectric is not situated in a constant but in a periodically varying electric field $E = E_0 \cos \omega t$, there flows through the condenser an A.C. which is proportional to $\mathrm{d}D/\mathrm{d}t$. When the dielectric is loss-free, the relation between D and E is here given by (1), so that

$$D = \varepsilon E_0 \cos \omega t = D_0 \cos \omega t \quad . \quad . \quad . \quad . \quad (2)$$

If, however, the insulator exhibits losses the following holds:

$$D = D_0 \cos (\omega t - \delta), \quad . \quad . \quad . \quad . \quad (3)$$

where δ is called the loss angle. The amplitude D_0 will in general differ from that in the ideal case (1), while both D_0 and δ may also still depend upon the frequency.

If a resistance R_p is connected in parallel with a condenser (capacity C), a phase difference δ occurs between current and voltage and

$$\tan \delta = \frac{1}{\omega C R_p}\,.$$

From this it may be seen that the occurrence of a phase shift is equivalent to the presence of a resistance in the circuit. As a result there is heat development in the dielectric, which is given per unit volume and per second by

$$W = {}^1\!/_2\, E_0 D_0 \omega \sin \delta \approx {}^1\!/_2\, E_0 D_0 \omega \tan \delta, \quad . \quad (4)$$

since usually $\delta < 0.01$ and thus the difference between δ, $\tan \delta$ and $\sin \delta$ can be disregarded.

This heat development may sometimes lead to disturbing rises in temperature.

Considering the importance in electrotechnics of having at one's disposal dielectrics which exhibit not only small dielectric losses but also a small temperature coefficient, it is understandable that these quantities have been measured for many substances. The results, as far as the losses are concerned, are usually characterized by giving the value of $\tan \delta$, in which δ is the phase shift from formula (3), while for the effective dielectric constant the quotient $D_0/E_0 = \varepsilon$ can be chosen [3]).

Results of the measurements

After these preliminaries we may now formulate the results of the measurements. It was found that many of the technical solid dielectrics investigated possess the following properties:

1) $\tan \delta$ is practically independent of the frequency in a wide region and entirely independent of the amplitude of the measuring field.
2) The dielectric constant ε a) is practically independent of the frequency and b) usually has a positive temperature coefficient.
3) Between $\tan \delta$ and the temperature coefficient of the dielectric constant ε there exists the relation

$$\frac{1}{\varepsilon}\frac{\partial \varepsilon}{\partial T} = A \tan \delta, \quad . \quad . \quad . \quad . \quad . \quad (5)$$

in which A depends only slightly on the nature of the substance and on the frequency and is inversely proportional to the absolute temperature T. At room temperature $A \approx 0{\cdot}05$.

In the table below several examples are given. The measurements were taken at room temperature and at a frequency of $1{\cdot}5 \times 10^6$ c/sec. It may be seen that, in spite of the fact that the values of $\tan \delta$ are very varied, the value of $A \times 10^2$ varies only between 4·4 and 9, and usually amounts to about 5.

Material	$\tan \delta \cdot 10^4$	$\frac{1}{\varepsilon}\frac{\partial \varepsilon}{\partial T} \cdot 10^6$	$A \cdot 10^2$
Quartz glass	1.0	9.0	9
Mica	1.1	5.5	5.0
Kersima *	8.1	51.0	6.3
Boron glass *	13.1	107	8.2
Cu-Zn-ferrite	32	141	4.4
Mycalex	38	190	5.0
G 40 glass *	40	240	6.0
Ebonite	81	460	5.7
Röntgen glass *	100	580	5.8
"Pliofilm"	300	1440	4.8
Celluloid	450	2600	5.8
"Pertinax"	550	3020	5.5
"Novotext"	850	4760	5.6

*) Products of Philips. "Kersima" is a ceramic product mainly Mg-silicate; G 40 is a NaK-borosilicate glass; Röntgen glass a NaCa-silicate glass.

The remarkable point about the materials having these properties is that they usually possess an irregular internal structure, *i.e.* the substances are in a vitreous state or at least exhibit a strong irregular disturbance of the single-crystal state.

It must be stated here that properties (1) and (2) had already been known for a long time. Already in 1914 Wagner had proposed a theory of the dielectric losses and especially of their variation with the frequency, from which the conclusion could be drawn that under special circumstances tan δ might be independent of the frequency in a wide region. Property (3), however, had not yet been observed, and in our opinion it is the most remarkable one. It also has important practical consequences. Search is often made for substances with a low temperature coefficient of the dielectric constant. The property in question here provides the evidence that in general such a search must not be conducted among the substances possessing a high value of the loss angle δ.

---

[3]) $E = E_0 \cos \omega t$ and the quantity D of formula (3) could also be written as complex quantities. The quotient $D/E = \varepsilon$ formed by analogy with (1) would then also become a complex number, of which $|\varepsilon| = D_0/E_0$ would represent the absolute value. In the notation of $|\varepsilon|$ for the quotient D_0/E_0 we have omitted the lines $||$ for the sake of simplification. *Cf.* also M. Gevers, Philips Res. Reports **1**, 197, 1946, especially p. 197 p. 198.

It is now our intention to show how not only the first but also the other properties can be made understandable by considerations which are in part very analogous to those of Wagner.

We assume, as is customary, that the dielectric losses in solid insulating substances such as the above are due to one of the following causes:

a) the presence of dipoles, which are only able to follow the alternations of an external field with a certain time lag;
b) the presence of separate semi-conducting regions in the substance, whereby the transport of charge to one side or other of the region exhibits an appreciable retardation;
c) the presence of free ions.

A combination of a), b) and c) can of course also occur. It has been found that each of the causes separately can bring about the occurrence of the properties found experimentally. For the present, however, we shall confine our considerations to the first case.

Explanation of the results of the measurements

The above-mentioned causes a), b) and c) have this in common, that they lead to after-effect phenomena.

In the article already referred to the influence of after-effect on the phenomena in a dielectric was discussed. It was found that if a constant field $E = E_0$ was suddenly $(t = 0)$ applied (*fig. 1*), the dielectric displacement behaves according to the formula:

$$\left.\begin{aligned} D &= E_0\,(\varepsilon_1 + \varepsilon_2 f(t)), \\ f(t) &= 1 - e^{-t/\tau} \text{ for } t > 0,\quad f(t) = 0 \text{ for } t < 0. \end{aligned}\right\} \quad (6)$$

The quantity ε is determined by the polarizability of the atoms; thus all atoms of the substance contribute more or less to ε_1. The quantity ε_2 is due to the mobile dipoles, so that usually $\varepsilon_2 \ll \varepsilon_1$. The quantity τ is called the relaxation time.

If, instead of a constant field, at the moment $t = 0$ a variable field $E = E_0 \cos \omega t$ is applied the expression for D becomes

$$D = K \cos \omega t + L \sin \omega t + M e^{-t/\tau}, \ldots\ldots \quad (7a)$$

where

$$K = E_0\,(\varepsilon_1 + \frac{\varepsilon_2}{1 + \tau^2\omega^2}),\ L = E_0 \varepsilon_2 \frac{\tau\omega}{1 + \tau^2\omega^2},\ M = \frac{-E_0\varepsilon_2}{1 + \tau^2\omega^2}. \quad (7b)$$

The third term in (7a) dies out quickly and for $t \gg \tau$ it may be ignored, so that (7a) is reduced to

$$D = K \cos \omega t + L \sin \omega t = \sqrt{K^2 + L^2} \cos(\alpha t - \text{arc tan}\,\frac{L}{K}), \quad (8)$$

which expression is of the form (3). Here $\varepsilon = \sqrt{K^2 + L^2}/E_0$ is thus the dielectric constant, while tan $\delta = L/K$.

If different relaxation times $\tau_2, \tau_3, \ldots$ are present simultaneously, then

$$K = E_0\,(\varepsilon_1 + \sum_m \frac{\varepsilon_m}{1 + \tau_m{}^2\omega^2}) \text{ and } L = E_0 \sum_m \frac{\varepsilon_m \tau_m \omega}{1 + \tau_m{}^2\omega^2}. \quad (9)$$
$$m = 2, 3, 4, \ldots$$

Since the values ε_m are usually very small compared with ε_1, L is negligible compared with K, and in many cases one may confine oneself to the first term in the expression for K.

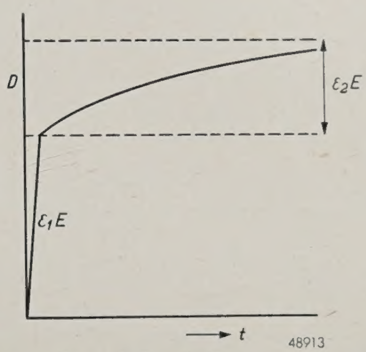

Fig. 1. The variation of the displacement D with the time t, when at time $t = 0$ an electric field E is suddenly switched on and causes the displacement $\varepsilon_1 E$. The contribution caused by the adjustment of the dipoles increases gradually from zero to $\varepsilon_2 E$.

For the explanation of the after-effects according to assumption (a) we now make use of considerations which were originally applied by Debije to liquid dielectrics. It is hereby assumed that in an insulating, viscous (thickly flowing) liquid there are a number of molecules, considered to be spherical, which are carriers of a dipole. In the first instance we suppose all molecules to be alike; let the number of them per cm³ equal n.

The formula which Debije arrives at and which of course shows great similarity to the previously [1]) derived expressions, is as follows:

$$D = E_0 \left(\varepsilon_1 + \frac{na}{1 + \tau^2\omega^2}\right) \cos \omega t + $$
$$+ E_0 \frac{na \cdot \tau\omega}{1 + \tau^2\omega^2} \sin \omega t, \ldots \quad (10)$$

where τ is the relaxation time and a the maximum possible contribution per particle to the dielectric constant. One thus finds in approximation for the dielectric constant and the phase angle

$$\varepsilon = \varepsilon_1 + \frac{na}{1 + \tau^2\omega^2}, \ldots\ldots \quad (11a)$$

$$\tan \delta = \frac{na}{\varepsilon_1} \frac{\tau\omega}{1 + \tau^2\omega^2} \ldots\ldots \quad (11b)$$

As a function of ω equation (11b) has a maximum at $\omega = 1/\tau$ (*fig. 2*).

The formulae, which were originally derived for a model of a liquid dielectric, are now also applied to the solid substance, although in this case we cannot form such a clear picture of the frictional forces opposing the orientation of the dipoles. It may be assumed that the dipoles present in the solid substance are bound to a number of preferential positions and that as a result, upon the transition from a given orientation to another, a number of "potential peaks" lying between the preferential positions must be exceeded at the expense of a certain energy q ("activation energy").

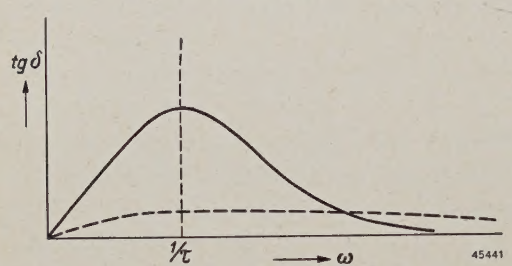

Fig. 2. The continuous line curve represents the variation of the tangent of the loss angle δ with the angular frequency ω according to Debije. τ is the time of adjustment (relaxation time). The dotted line curve relates to the case where different groups of dipoles are present with different values of τ.

It is important for our argument that τ is very closely dependent on the temperature. From the kinetic theory of heat it follows that

$$\tau = \tau_0\, e^{q/kT},$$

where τ_0 has the significance of a material constant. Further q/k, where k is the Boltzmann constant, is of the order of magnitude 10 000 °K.

Theory as well as experiment show that τ_0 is a maximum of 10^{-13} sec but may also be considerably less, for instance 10^{-20} sec.

After these preliminaries on the calculations of Debije and the relaxation time τ — both of which related to insulators with a very regular internal structure (all dipoles alike) — we now return to the technical dielectrics.

The special assumption from which the three above-mentioned properties follow is that in the technical solid dielectrics investigated, which all possess an irregular internal structure, there is a wide scattering in the values of the activation energy q. This means, therefore, that not all the dipoles possess the same relaxation time and that the differences which occur are caused by differences in the values of q. All values of q may occur in a certain interval.

We are most interested in property (3) and the constant A thereby occurring. Theory based on the special assumption just mentioned gives the following formula:

$$A = -\frac{2}{\pi}\,\frac{\ln \tau_0\omega}{T} \quad . \; . \; . \; . \; . \; . \quad (12)$$

We shall here show briefly how the three properties (1), (2) and (3) follow from the fundamental assumption indicated and how in particular the formula for the constant A is derived. To this end we must first discuss the assumption in more detail. Let n be the total number of dipoles present per cm³. We then assume that the part of this number which upon being placed in an electric field shows a value of q lying between q and $q + dq$ can be represented by

$$n \cdot G\,(q)\, dq \qquad (\text{with } \int G\,(q)\, dq = 1),$$

where $G(q)$ in a given region varies only slowly with q. We shall be able to formulate this last condition more precisely in the course of our further considerations. The variation of $G(q)$ may, however, differ considerably between one substance and another. Thus per cm³ we encounter in the substance: $n \cdot G(q)dq$ dipoles all possessing the relaxation time $\tau = \tau_0 e^{q/kT}$.

For the sake of simplification we assume that τ_0 and a do not depend upon q. It is, however, easy to ascertain that the method of calculation to be followed also remains valid when these quantities do indeed depend upon q, but not very closely.

For the total dielectric constant and the phase angle we now find

$$\varepsilon = \varepsilon_1 + na \int_0^\infty \frac{1}{1 + \tau^2\omega^2}\, G(q)\, dq \; . \; . \; . \quad (13a)$$

$$\text{and } \tan\delta = \frac{na}{\varepsilon_1} \int_0^\infty \frac{\tau\omega}{1 + \tau^2\omega^2}\, G(q)\, dq\,, \; . \; . \; . \; . \; . \quad (13b)$$

where $\tau = \tau_0 e^{q/kT}$, while ε_1 represents the ordinary dielectric constant which the substance would exhibit if the dipoles were unable to move.

From these expressions, which are nothing but a generalization of formulae (11a) and (11b) of Debije, the three properties in question now follow.

We shall, however, first reduce these formulae somewhat in order to make it easier to draw conclusions from them. For ε one may write

$$\varepsilon = \varepsilon_1 + na \int_0^{q_m} G\,(q)\, dq, \; . \; . \; . \; . \quad (14a)$$

where q_m is the value of q for which $\tau\omega = 1$, namely $q_m = -kT \ln \tau_0\omega$, while for $\tan\delta$ we may write

$$\tan\delta = \frac{\pi}{2}\,\frac{na}{\varepsilon_1}\, G(q_m)\, kT \; . \; . \; . \; . \; . \quad (14b)$$

Formulae (14a) and (14b) may be interpreted as follows: Only those dipoles contribute to ε whose relaxation time is smaller than $1/\omega$ (the others are too "slow"); the losses are caused by a group of dipoles whose relaxation time amounts to about $1/\omega$.

In the derivation of (14a) and (14b) use has been made of the property that $G(q)$ changes "slowly" with q. The requirement is that G shall vary little in an interval for q in which $\tau\omega/1 + \tau^2\omega^2$ differs appreciably from zero.

From the expressions (14a) and (14b) for ε and $\tan\delta$ the peculiarities found experimentally now follow, and we shall examine these successively for the properties (1), (2) and (3).

Property (1)

In order to understand that tan δ is practically independent of ω we note that $q_m = -kT \ln \tau_0\omega$ shows only slight variations with ω. Because, when $\tau_0 = 10^{-14}$ sec, then

$$-\ln \tau_0\omega = 27{\cdot}6 \qquad \text{at } \omega = 10^2 \text{ rad/sec},$$
$$-\ln \tau_0\omega = 11{\cdot}5 \qquad \text{at } \omega = 10^9 \text{ rad/sec}.$$

The change in q_m is thus only by a factor 2·5 upon the enormous change in ω by a factor 10^7. At smaller values of τ_0 the variation is still smaller. Now because G depends little on q_m, in a suitable region tan δ will depend little on ω. The "suitable region" need thus occupy at the most a factor 2·5 in the values of q in the vicinity of q_m.

A dependence of tan δ on the amplitude E_0 does not occur at all.

Properties (2a) *and* (2b)

In order to prove the property (2a) it must be recalled that

$$\frac{\partial\varepsilon}{\partial\omega} = \frac{\partial\varepsilon}{\partial q_m} \cdot \frac{dq_m}{d\omega} = -\,G(q_m)\frac{kT}{\omega} = -\frac{2\varepsilon_1}{\pi\omega}\tan\delta.$$

In the frequency region where tan δ is practically independent of ω, therefore, as may be found by integration of the above formula, the following is valid:

$$\varepsilon - \varepsilon(\omega_0) = -\frac{2\varepsilon_1}{\pi}\tan\ln\delta\,\frac{\omega}{\omega_0},$$

where ω_0 is the beginning of the frequency region in question. For the sake of example we choose $\omega_0 = 10^2$ rad/sec and $\omega = 10^9$ rad/sec, while for instance we let tan $\delta = 10^{-2}$ at $\omega = \omega_0$. Then $\varepsilon = \varepsilon(\omega_0) - 2\varepsilon_1/\pi$ tan δ ln 10^7. Since tan $\delta = 10^{-2}$, we obtain $\varepsilon = \varepsilon(\omega_0) \cdot (1-0{\cdot}1)$. Upon a change in frequency from 10^2 to 10^9 rad/sec therefore ε changes only by 10%.

From (14a) it follows that

$$\frac{\partial\varepsilon}{\partial T} = \frac{\partial\varepsilon_1}{\partial T} + \frac{\partial\varepsilon}{\partial q_m} \cdot \frac{dq_m}{dT} = \frac{\partial\varepsilon_1}{\partial T} - na\,G(q_m)\,k\ln\tau_0\omega. \qquad (15)$$

In general it is found that $\partial\varepsilon_1/\partial T$, caused by the expansion of the substance upon rise in temperature, is negligible compared with the second term.

At the values of τ and ω occurring in practice the second term is always positive. We therefore find that the temperature coefficient is positive, while in a good approximation the following is valid:

$$\frac{\partial\varepsilon}{\partial T} = -na\,G\,(q_m)\,k\ln\tau_0\omega \quad . \; . \; . \; . \; . \quad (16)$$

Property (3)

In order to derive the remarkable property $(1/\varepsilon)\,\partial\varepsilon/\partial T = A \tan\delta$ we note that formula (16), just found by making use of the expression for tan δ, can be written as follows:

$$\frac{\partial\varepsilon}{\partial T} = -\frac{2\,\varepsilon_1\tan\delta}{\pi T}\ln\tau_0\omega,$$

so that

$$\frac{1}{\varepsilon}\frac{\partial\varepsilon}{\partial T} = -\frac{2}{\pi}\frac{\ln\tau_0\omega}{T}\tan\delta = A\tan\delta. \; . \; . \quad (17)$$

Thus

$$A = -\frac{2}{\pi}\frac{\ln\tau_0\omega}{T} \quad . \; . \; . \; . \; . \; . \; . \quad (12)$$

The "constant" A thus determined depends only very little on the frequency, because if we choose $T = 300°$ K, at $\tau_0 = 10^{-14}$ sec, we find

$$A = 0.058 \quad \text{at } \omega = 10^2 \text{ rad/sec and}$$
$$A = 0.024 \quad \text{at } \omega = 10^9 \text{ rad/sec},$$

while at $\tau_0 = 10^{-20}$ sec

$$A = 0.088 \quad \text{at } \omega = 10^2 \text{ rad/sec},$$
$$A = 0.054 \quad \text{at } \omega = 10^9 \text{ rad/sec}.$$

From this it follows that, although the values of τ_0 for the different substances differ by many times a factor of 10, at room temperature we shall always find only slightly differing values of A for measurements in the customary frequency region. The calculated value of A agrees reasonably well with the measured values which lie in the vicinity of 0·05.

Until now we have specialized on the case (a) where the losses are supposed to be caused by dipoles. We may now give an entirely similar line of reasoning for cases (b) and (c) in which semi-conducting regions or free ions are the cause. It would lead us too far to treat this possibility in as great detail as has been done for case (a). The fact that the properties found can be derived not only from cause (a) but also from (b) and (c), however, also implies that for a given substance it is difficult to ascertain whether the losses are caused by (a), by (b) or by (c) or by more than one of these together.

Exceptions

In the above explanation of property (3) it was assumed that $\partial\varepsilon_1/\partial T$ is small enough to be ignored. In most cases this is true, but there are a few exceptions. Theory shows (see footnote [2]) that for the part ε_1 of the dielectric constant ε which is independent of the after-effect the following equation holds:

$$\frac{1}{\varepsilon_1}\frac{\partial\varepsilon_1}{\partial T} = -\,\alpha_{\text{lin}}\frac{(\varepsilon_1-1)\,(\varepsilon_1+2)}{\varepsilon_1}, \qquad (18)$$

where α_{lin} represents the linear coefficient of expansion. From this it follows that $\partial\varepsilon_1/\partial T$ is negative and in absolute value larger according as α_{lin} and ε_1 are larger.

Now while on the one hand the coefficient of expansion seldom assumes abnormally large values, high values of ε_1 (and thus also of ε) do on the other hand certainly occur. In contrast to most substances having a value of ε between 1 and 10, for the substance rutile (TiO_2), for example, $\varepsilon \approx 100$. For such substances $d\varepsilon_1/dT$ can no longer be ignored, as was assumed in the derivation of the formula for A (formula (16)). As a result properties (2b) and (3) do not hold for these substances. Under the influence of the term $\partial\varepsilon_1/\partial T$ they have a lower positive temperature coefficient than would be expected according to formula (5), and in some cases, especially when tan δ is small, they even have a negative temperature coefficient.

In cases where a substance is needed with a very small temperature coefficient of the dielectric constant, a choice can thus be made in two ways. If it is desired to use substances with normal values of ε, for which property (3) holds, as was stated at the beginning, the search must be made among the substances possessing a low value of the loss angle δ. The other possibility is that one should use substances with a high value of ε for which property (3) does not hold. With these substances it is possible, if desired, to obtain negative temperature coefficients.

Number 1 of Volume 2 (February, 1947) of *Philips Research Reports* contains the following papers:

R 32: B. D. H. Tellegen: Coupled circuits.

R 33: W. Elenbaas: On the excitation temperature, the gas temperature, and the electron temperature in the high-pressure mercury discharge.

R 34: H. B. G. Casimir: On the theory of eddy currents in ferromagnetic materials.

R 35: H. C. Hamaker: Radiation and heat conduction in high-scattering material.

R 36: H. A. Klasens: The light emission from fluorescent screens irradiated by X-rays.

Readers interested in any of the above mentioned articles may apply to the Administration of the Philips Physical Laboratory, Kastanjelaan, Eindhoven, where a limited number of copies are available for distribution. For a subscription to Philips Research Reports please write to the publishers of Philips Technical Review.